Genetic Susceptibility to Traffic Related Pollutants

Jamaludin, Jeenath Banu

Awarding institution:
King's College London

GENETIC SUSCEPTIBILITY TO TRAFFIC RELATED POLLUTANTS

Jeenath Banu Jamaludin

Analytical & Environmental Sciences Division
School of Biomedical & Health Sciences
King's College London

Thesis submitted to the King's College of London
in fulfilment of the requirements for the
degree of Doctor of Philosophy

November 2014

Declaration:

I, Jeenath Banu Jamaludin declare that all the work submitted in this thesis is my own, except where stated otherwise.

Signed:_____ (Student)

Signed:_____ (Supervisor)

Date: 30th June 2015

Abstract

A strong correlation exists between acute and chronic exposure to traffic derived pollutants and poor respiratory health. Specifically, diesel exhaust (DE) components such as NO_2 and fine particles ($PM_{2.5}$) have been related to impaired lung growth and increased respiratory and allergic symptoms in children and adults living near busy roads. On this basis, implementation of strategies to reduce diesel emissions and improve air quality should provide a measureable improvement in the respiratory health of populations resident in high traffic areas. The introduction of the London Low Emission Zone (LEZ), the largest of its kind in the world, covering an area of 2,644 km^2 and a resident population of more than 8 million, provided a unique opportunity to examine this, as well to quantify the impact of DE emissions on the respiratory health of London's population.

London's Low Emission Zone was introduced as part of the Mayor of London's Air Quality Strategy, with the aim of improving public health through targeted reductions in tail pipe emissions from the most polluting vehicles entering the city. The objective of decreasing PM_{10} concentrations was to be achieved by restricting the entry of the oldest and most polluting diesel vehicles (heavy goods vehicles (HGVs), buses and coaches, larger vans and minibuses) into Greater London by providing incentives to operators to upgrade their fleets to lower emission vehicles. The Low Emission Zone was enacted as a phased tightening of emission standards for each vehicle class, with the first phase coming into force at the beginning of February 2008. This applied to HGVs greater than 12 tonnes and restricted entry to the zone for those vehicles not meeting the Euro III emissions standard for PM_{10}. Phase 2 followed in July 2008 widening restrictions to include HGVs between 3.5 and 12 tonnes, buses and coaches. Vehicles failing to meet these emissions standards within the zone were initially charged £200 (£100 for vans an minibuses) per day, with enforcement achieve using cameras to identify the registration numbers of vehicles and the Driver and Vehicle Licensing Agency (DVLA) database to identify a vehicle's emissions standard. In its initial configuration phase 3, restricting access to heavier LGVs and mini-buses not meeting Euro III PM standard was planned for October 2008.

In our initial four year study design we planned to examine the respiratory health of cross sectional panels of 8-9 year old school children living within the zone from November

2008; encompassing the first two years post phases 1 and 2, and two years post phase 3 (see **Figure s1**). As the subject recruitment and health assessments began in November 2008, this afforded us the opportunity of addressing the impact of the third phase of the LEZ, by comparing lung function and respiratory symptoms in the two years before and after phase 3. We also planned to examine year-on-year changes related to projected reductions in vehicle emissions as newer cleaner vehicles entered the fleet, independent of the LEZ, and the increased period the children had lived within the zone, from 11-15 months (11.5-13.9% of lifetime) in year 1, to 44-60 months (45.8-55.6%) in year 4. During annual school visits, children were asked to perform spirometry and provide a urine sample for the assessment of exposure (metals, reflective of defined traffic sources) and response biomarkers (oxidative damage). In addition, the parents/guardians of the children completed a questionnaire on respiratory / allergic symptoms and the children provided DNA samples to investigate genetic susceptibility to the detrimental effects of air pollution, focusing on a panel of antioxidant and xenobiotic genes, as well as a genetic marker associated with the onset of childhood asthma.

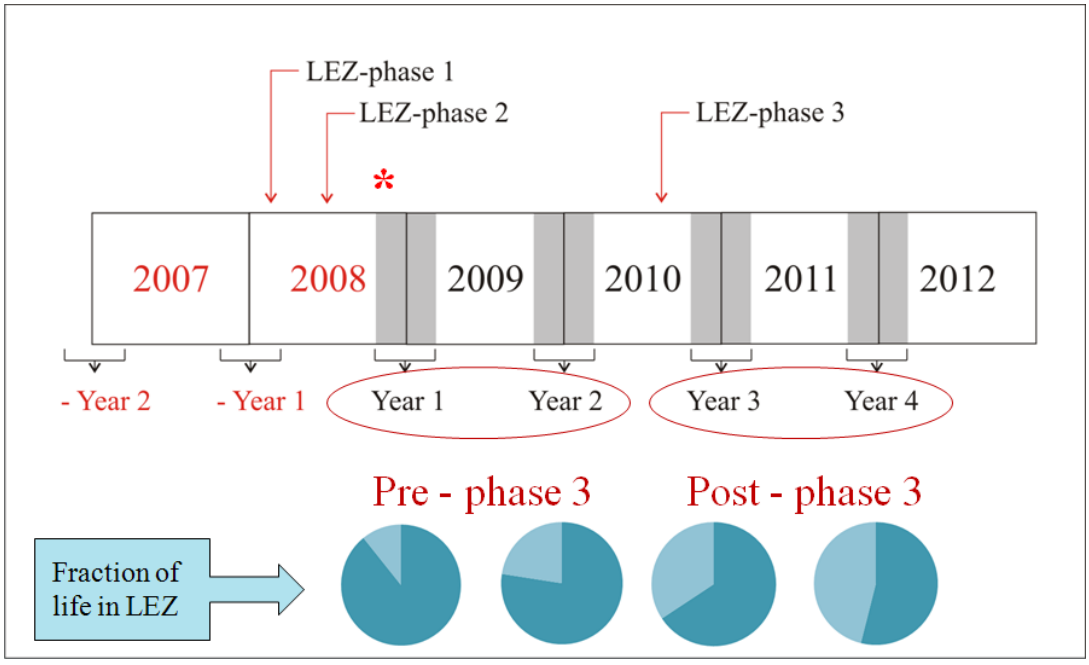


Figure s1 The original LEZ study design outlining the winter sampling periods (shaded gray) across the four year study with the introduction of phase 3 of the scheme permitting a two year pre- and post-assessment. The implementation dates of phase 1 and 2 restrictions are illustrated as in the increased proportion (light blue area) of the children lifetime spent resident within the zone over the study period.

In May 2008 Boris Johnson was elected the new Mayor of London, with a manifesto commitment to review ongoing traffic management schemes within the city, including the LEZ, and on the 2nd February 2009 he announced intention to cancel the third phase of the LEZ, subject to the outcome of a public consultation later in the year. This political decision therefore robbed us of the original intervention we were planning to address in our original design. Following a further consultation, the scheme was finally fully implemented and expanded on the 3rd of January 2012 (LEZ phase 3 and 4), with Euro III emission standards for minibuses and vans and a further tightening of emission standards (Euro IV) on Lorries over 12 tonnes, between 3.5-12 tonnes, as well as buses and coaches. In light of this development we obtained additional funding to examine furthers panels of school children in Nov 2012 - March 2013 and Nov 2013 - March 2014, extending our study to six years, allowing a formal assessment of the three years pre and two years post LEZ phase 3 and 4, with year 4 straddling the periods of phase 3 and 4 implementation (**Figure s2**). Children at the conclusion of the study in March 2014, who have been resident within the LEZ since birth will have lived within the zone for 68.8-83.3% of their lives.

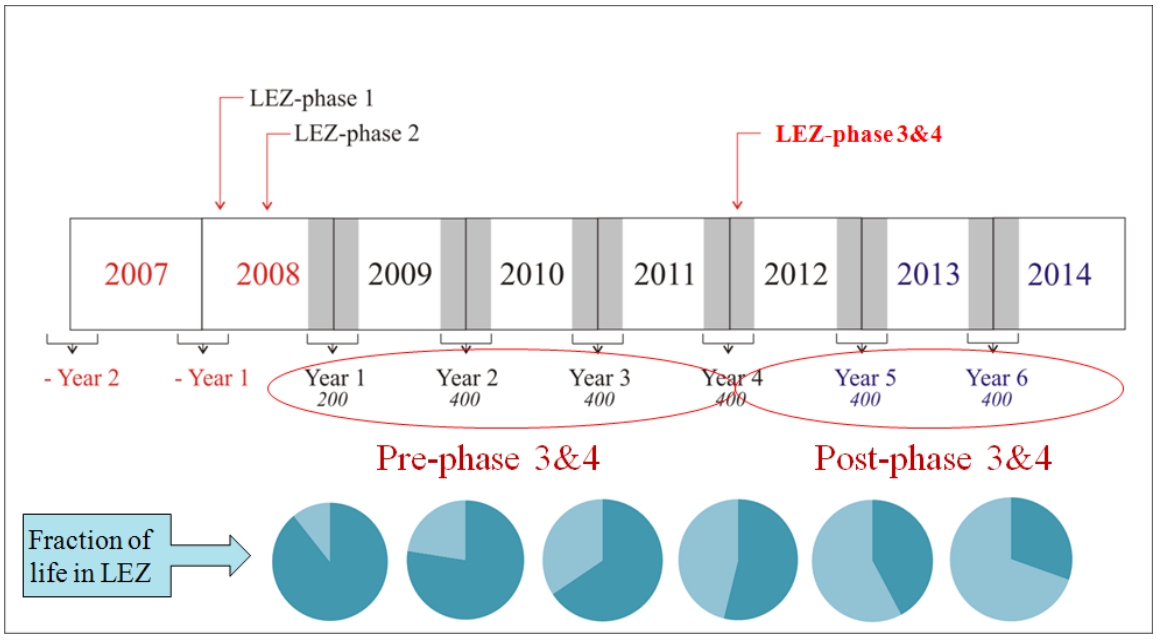


Figure s2 The revised LEZ study design outlining the winter sampling periods across the six year study with the introduction of phase 3 of the scheme permitting a two year pre- and post-assessment.

The data presented in this thesis is therefore based upon the first three years of the study and therefore constitutes a baseline analysis of the relationship between air pollution in London and our key respiratory endpoints prior to the formal evaluation of Phase 3 and 4 in 2014/15.

In the first experimental chapter (**Chapter 3**) I evaluated the associations between traffic-related air pollutants and respiratory/allergic symptoms within our cross-sectional children's cohort. Information on respiratory/allergic symptoms was obtained using a parent-completed questionnaire and linked to modelled annual air pollutant concentrations based on the residential address of each child, using a multivariable mixed effects logistic regression analysis. Exposure to traffic-related air pollutants was associated with current rhinitis (NOx [OR 1.01, 95% CI 1.00-1.02], NO_2 [1.03, 1.00-1.06], PM_{10} [1.16, 1.04-1.28] and $PM_{2.5}$ [1.38, 1.08-1.78], all per $\mu g/m^3$), but not with other respiratory/allergic symptoms. Furthermore, over the first three years of the operation of London's LEZ I did not observe evidence of reduced ambient air pollution levels, or year-on-year changes in the prevalence of respiratory/allergic symptoms. I found no evidence that these associations were modified by polymorphisms in gasdermin B, located at the chromosome 17q12, associated with the risk of childhood asthma. These data confirm previously reported associations between traffic-related air pollutant exposures and symptoms of current rhinitis. Importantly, whilst the data is largely confirmatory, this remains one of the few studies that has addressed respiratory symptoms in urban children over the period of rapid dieselization within Europe.

In **Chapter 4** I report evidence of reduced lung volumes (FVC - Forced Vital Capacity) in children living within the study area. This negative association was small and most strongly associated with modeled annual NOx concentrations, at the residential address level. A straightforward method to discriminate between acute versus chronic pollutant effects was developed for the study. Acute exposures were assessed by scaling annual mean concentrations according to a 'Nowcast' factor calculated for each pollutant for the period immediately prior to the health assessment. This factor was defined as the ratio between concentrations measured by a local subset of London Air Quality Network monitoring sites in the prior period, and the annual mean measured by the same sites. Using this approach I was able to dissect out whether basal lung function was related to short or long term exposures. In the absence of relationships between FEV_1 and FVC with 24 hour and 7 day average exposures, the association between FVC and annual pollutant exposures was interpreted as reflecting evidence of impaired lung growth. In this initial analysis I found no evidence that polymorphisms in the commonly studied glutathione S transferases (GSTM1 and GSTP1) and NADPH quinone oxidoreducatse

(NQO1) genes modified the association between lung function and pollutant exposure. In a secondary analysis I examined whether polymorphisms in Cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) and the Aryl hydrocarbon receptor (AhR) might modify the association between pollutant exposures and lung function, based on their role in the xenobiotic metabolism of Polycyclic aromatic hydrocarbons (PAHs). This is the first time polymorphisms in these genes have been investigated in the context of air pollution – lung function interactions. I found evidence that two SNPs in CYP1A1 influenced acute lung function responses to air pollution, one being protective and the other detrimental. Again all the significant relationships were with FVC and not FEV_1, in contrast to much of the pre-existing literature. No noticeable improvement in respiratory health was observed in the children studied over the first three years of the LEZ.

In the final results chapter (**5**) I explored the relationship between pollutant exposures in the same cohort of children with urinary biomarkers of oxidative stress: 8-oxo-2'-deoxyguanosine (8-oxodG, as a marker of DNA damage) and 8-isoprostane (lipid oxidation). In addition, I also explored whether the determination of a panel of urinary metals would be informative as a biomarker of traffic exposure. These data demonstrated a positive relationship between urinary 8-oxodG concentrations and short term, acute particulate exposures (both PM_{10} and $PM_{2.5}$). These responses appeared to reverse with longer term exposures (both 7 day and annual), i.e. high pollutant exposures were associated with lower urinary 8-oxodG concentrations, suggesting some form of biologic adaption to the pollutant stress. I therefore examined a panel of SNPs in genes, either directly related to oxidative stress (GCLM, Glutamate-cysteine ligase regulatory subunit and SOD3, extracellular superoxide dismutase), or the coordination of the antioxidant response (Nrf2, Nuclear factor (erythroid-derived 2)-like 2). Whilst polymorphisms in SOD3 did not modify the relationship between urinary markers of oxidative stress and the pollutant exposures, there was evidence that SNPs in Nrf2 appeared to blunt the sub-chronic exposure adaption. The work on the use of urinary metals as biomarkers did not prove fruitful. Therefore overall this series of studies have demonstrated long term effects of primary pollutants on indices of lung function and systemic oxidative stress.

Acknowledgements

It is a great pleasure to acknowledge and thank everyone who helped and inspired me towards completing my thesis over the past 4 years.

I would like to offer exceptional thanks to my supervisor Dr. Ian Mudway, for his valuable support, his professional guidance and knowledge. Especially his constructive ideas, valuable critiques and assistance, which helped me to shape and accomplish this PhD and thesis writing. Your motivation and precious time has been an inspiration throughout my study and writing. I would like to extend my sincere gratitude to Prof. Frank Kelly, my co-supervisor to facilitate the opportunity to use his lab for my research and his valuable scientific advice. It has been a great privilege working with you.

Not forgetting the ever supportive other members of the team including Prof. Chris Griffith, Prof. Robert Walton, Dr. Isobel Dundas, and Dr. Stephen Bremner. Many thanks to Helen Wood who have been keeping the master database tidy and up to date. I cannot thank enough to Nadine Marlin, the statistician, who had worked tirelessly in analyzing the results for the study. My gratitude to the rest of the Lung Biology Group: Chrissi Dunster, Melis Bicer, Irene Katsaiti and Esme Purdie for sharing the good and bad times in our research experience. Not forgetting all the past and present associates of my laboratory for their various roles and support in completing my thesis.

I would like to express great appreciation to my sponsors, University Science of Malaysia, particularly Institute For Research in Molecular Medicine (INFORMM) and the Ministry of Higher Education of Malaysia (MOHE). This achievement was not possible without significant financial support provided by you. Thank You!

Thanks to the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London, for providing the funding for this research. Over the years during my PhD, I have been privileged to work with a number of school children and their school teachers, who allowed me to collect the data for the study. Given the nature of this thesis, no devotion would be complete without a compliment to their selfless time, tolerance and determination.

My deepest and humble gratitude goes to my mother, Noorjahan, my dad, Jamaludin and my sisters, Anis and Nurul and brother Abbas, for their prayers, constant encouragement, inspiration and guidance throughout my writing. I would like to dedicate this thesis to my husband, Khalid Ahmad Sheikh, for his emotional support, who has been the pillar towards completing this thesis. Who saw me through mental breakdown with time and care when writing seemed so impossible. Also to my ever loving children Faris, Farah Dina and Farhan, who often put a big smile on my face when I find myself buried in what seems never ending writing. I am eternally thankful to my nanny, Fatimah, who has dedicated her life taking care of my children with patience and affection. There are no words I could describe here to which my heart feels for all your love and support, which has enabled me to bring this work to completion. Thank You! All praise is due to Allah.

Table of Contents

Chapter 2: General Materials and Methods

Chapter 3: Air Pollution, the London Low Emission Zone, and the Prevalence of Respiratory and Allergic Symptoms in Schoolchildren in East London

Appendices

List of Figures

List of Tables

List of Commonly Used Abbreviations

µg/m^3	microgram per cubic meter
1-OH-P	1-hydroxypyrene
8-OHdG	8-hydroxy-2-deoxyguanosine
8-oxodG	8-Oxo-2'-deoxyguanosine
A	Adenine
AA	ascorbate
A230	absorbance at 230nm
A260	absorbance at 260nm
A280	absorbance at 280nm
ADBR2	beta 2-adrenergic receptor
AhR	aryl hydrocarbon receptor
Al	aluminium
Ala	Alanine
AhR	aryl hydrocarbon receptor
AP-1	activator protein 1
ARE	antioxidant responsive element
Arg	Arginine
ARG	arginase
As	arsenic
ASO	allelic specific oligonucliotide
Asp	aspartic acid
Asn	Asparagine
Ba	Barium
B[a]P	benzo[a]pyrene

BAL	bronchoalveolar lavage
BAMSE	the Children, Allergy, Milieu, Stockholm, Epidemiological Survey
BMI	body mass index
C	cytosine
CC16	Clara cell 16kDa protein
CCR	creatinine
Cd	cadmium
CHS	Children Health Cohort
Chr	chromosome
Cr	chromium
CO	carbon monoxide
COPD	chronic obstructive pulmonary disease
c-PAH	carcinogenic polycyclic aromatic hydrocarbons
Cr	chromium
cSNP	coding SNP
Cu	copper
CYP1A1	cytochrome P-450 enzymes
CYP1A2	cytochrome P450 1A2
CYP1B1	cytochrome P450 1B1
CYP2E1	cytochrome P450 2E1
CYPs	cytochrome P450s
Cys	Cysteine
dbSNP	database of Single Nucleotide Polymorphisms
DCM	dichloromethane
DE	diesel exhaust
DEP	diesel exhaust particles
DNA	Deoxyribonucleic acid
EBC	exhaled breath condensate

EC-SOD	extracellular superoxide dismutase
EDTA	ethylenediaminetetraacetic acid
EPHX1	microsomal epoxide hydrolase 1
ETS	environmental tobacco smoke
FAD	flavin adenine nucleotide
Fe	iron
Fe^{+2}	ferrous ion
Fe^{+3}	ferric ion
FEF	forced expiratory flow
FEF_{25-75}	forced expiratory flow between 25% and 75%
FeNO	fractional of exhaled nitric oxide
FEV	forced expiratory volume
FEV_1	FEV in 1 second
FPG	formamidopyrimidine DNA glycosylase
FVC	forced vital capacity
G	Guanine
GC	GenCall
GCLC	glutamate cysteine ligase catalytic subunit
GCLM	glutamate cysteine ligase modifier subunit
GIS	geographic information system
Gln	Glutamine
Gly	Glycine
GPx	glutathione peroxidase
GSH	reduced glutathione
GSR	glutathione reductase
GSS	GSH synthetase
GSSG	glutathione disulphide
GST	glutathione S-transferase

GSTA1	glutathione S-transferase A1
GSTM1	glutathione S-transferase mu
GSTP1	glutathione S-transferase pi
GSTT1	glutathione S-transferase theta
H_2O_2	hydrogen peroxide
His	histidine
HLA	human leukocyte antigen
HMOX1	heme oxygenase 1
HO-1	heme oxygenase 1
hOGG1	8-oxoguanine DNA glycosylase
HWE	Hardy Weinberg equilibrium
IgE	immunoglobulin E
IL-10	interleukin 10
IL-6	interleukin 6
IL-8	interleukin 8
Ile	isoleucine
Ile105Val	isoleucine 105 valine
IMD	index of multiple deprivation
iNOS	inducible nitric oxide synthase
LD	linkage disequilibrium
Leu	leucine
LEZ	Low Emission Zone
LSO	locus-specific oligo
LTA	lymphotoxin
MAF	minor allele frequency
mL	milli liter
mM	milli molar
Mn	manganese

MnSOD	manganese superoxide dismutase
Mo	molybdenum
MRE	metal responsive elements
MS	methionine synthase
MTHFR	methylenetetrahydrofolate reductase
NADH	nicotinamide adenine dinucleotide reduced
NADP	nicotinamide adenine dinucleotide phosphate
Ni	nickel
NF-KB	nuclear factor kappa-light-chain-enhancer of activated B cells
nm	nanometer
NO_2	nitrogen dioxide
NOx	nitrogen oxides (NO + NO_2)
NQO1	NADPH quinone oxidoreductase 1
Nrf2	nuclear factor erythroid 2 related factor 2
nsSNP	non-synonymous single nucleotide polymorphism
O_2•	superoxide radical
O_3	ozone
OH•	hydroxyl radical
OPA	oligo pool assay
PAH	polycyclic aromatic hydrocarbon
PCR	polymerase chain reaction
PCR-RFLP	PCR-restriction fragments length polymorphism
PEF	peak expiratory flow
PM	particulate matter
$PM_{0.1}$	particulate matter with aerodynamic diameter < 0.1μm
PM_{10}	particulate matter with aerodynamic diameter < 10 microns
$PM_{2.5}$	particulate matter with aerodynamic diameter < 2.5 microns
Pos	positive
ppb	parts per billion

ppm	parts per million
Pro	Proline
pSNPs	protein single nucleotide polymorphism
QC	quality control
qPCR	quantitative polymerase chain reaction
ROS	reactive oxygen species
rpm	revolution per minute
rs	reference SNP
rSNP	regulatory single nucleotide polymorphism
Sb	antimony
SB	strand breaks
SD	standard deviation
Ser	Serine
Sn	tin
SNP	single nucleotide polymorphism
SO_2	sulfur dioxide
SOD	superoxide dismutase
SP	saliva spit
S-PMA	S-phenylmercapturic acid
SS	saliva sponges
T	thymine
TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TE	tris EDTA
TGF-ß1	transforming growth factor beta 1
Thr	Threonine
Ti	titanium
TLR4	toll-like receptor 4
TNF-α	tumor necrosis factor alpha

Tm	melting temperatures
Tyr	Tyrosine
UA	urate
UFP	ultrafine particulate matter
UK	United Kingdom
UV	ultraviolet
V	vanadium
Val	Valine
W	tungsten
WGA	whole genome amplification
WHO	World Health Organization
XME	xenobiotic metabolizing enzymes
XPD	xeroderma pigmentosum D
XRCC1	X-ray repair cross-complementing group 1
XRE	xenobiotic responsive elements
Zn	zinc
ZnSOD	zinc superoxide dismutase

Chapter 1

Introduction

1.1 Health Effects of Environmental Pollutants

Long term exposure to air pollution, either within indoor or outdoor environments, has been shown to be associated with serious adverse health effects in both healthy and compromised individuals. Short-term exposure to air pollution has been shown to result in increased mortality in sensitive sub-populations and is associated with increased hospital visits due to acute effects such as irritation of the nose and eyes, headache, fatigue, wheezing, and an increased susceptibility to respiratory infections. In contrast, long-term exposure may result in impaired lung growth, remodelling of the airway and increased susceptibility of individuals to chronic lung diseases such as asthma, or chronic obstructive pulmonary disease (COPD) (WHO 2004).

1.1.1 Impact on Mortality and Morbidity

1.1.1.1 Epidemiological Studies

A large number of studies have been published reporting increased rates of pulmonary mortality and morbidity associated with exposure and elevated concentrations of ambient particulate matter (PM). These studies can be sub-divided into those addressing short- and long-term exposures in the general population and selected cohorts. Although studies focusing on short-term exposures of PM predominate, research on long-term exposures appears to be of greater importance. One such study which resulted in the establishment of the current United States ambient air

quality standards (Krewski et al. 2005), was conducted by Dockery et al. (1993) which focused on the relationship between air pollution in six cities in the USA with adult mortality. This cohort study covered a period of 14-16 years during which the mortality of 8,111 adults aged 25 to 74 was followed and related to the pollutant concentrations in each of the six cities. The pollutants measured included ozone, sulphur dioxide, suspended sulphate particles and total $PM_{2.5}$ (particles with an median aerodynamic diameter of less than 2.5μm). Their study revealed significantly increased rates of pre-mature death in the cities with the highest pollutant concentrations with the strongest associations observed with sulphate particles and $PM_{2.5}$, with evidence that the excess deaths were attributable to cardiopulmonary causes. In one recent multicity study, exposure to elevated $PM_{2.5}$ concentrations was associated with a significant increased risk of mortality from cardiovascular disease, respiratory disease, myocardial infarction and stroke (Zanobetti and Schwartz 2009).

In all of these studies, the health effects are related to the ambient concentration of PM, which is insensitive to regional variations in composition and hence source. Evidence that PM arising from certain sources may cause a specific risk is supported by studies examining the relationship between respiratory health and the proximity of populations to traffic (Hoek et al. 2002; Finkelstein et al. 2004), particularly in susceptible groups such as children and the elderly (Bobak and Leon 1992; Smargiassi et al. 2006; Woodruff et al. 2006). Janssen et al. (2003) measured the short-term effects of exposure to heavy-duty traffic-derived $PM_{2.5}$ and NO_2 on the respiratory health of children (7–-12 years) attending schools within 400 meters of busy motorways in the Netherlands. Over their study period (April 1997 – May 1998), respiratory and allergic symptoms, including persistent asthma, bronchitis and eczema were all significantly increased in children attending schools near motorways carrying a high percentage of heavy goods (diesel powered) vehicles. Children with pre-existing pulmonary or allergic conditions exhibited a worsening of their symptoms, which was related to their schools' proximity to motorways with high truck counts.

1.1.1.2 Human Exposure Studies

Whilst epidemiological studies provide evidence on the statistical association between air pollutant exposures and indices of health, they are unable to provide information on the mechanistic basis for the effects observed. With regard to exposure to vehicle derived pollutants healthy subjects exposed to diesel exhaust PM_{10} (100 and 300 $\mu g/m^3$) with associated gases, under controlled chamber conditions have been shown to exhibit transient increases in neutrophils, mast cells and T lymphocytes (Behndig et al. 2006; Salvi et al. 1999). This inflammatory response has been shown to be associated with alterations in the antioxidant composition of the airway lining fluids and an up-regulation/nuclear mobilization of redox sensitive transcription factors (Behndig et al. 2006; Pourazar et al. 2005). Exposure to a high dose of diesel (300 $\mu g/m^3$ for 1 hour) has also been shown to induce a delayed (24h post-diesel exposure) airway hyper-responsiveness to metacholine (Nordenhäll et al. 2001). This has been speculated to reflect an acute up-regulation of IL-10 in the bronchial mucosa (Holgate et al. 2003). A recent real-world experiment that examined the respiratory responses of asthmatics walking along Oxford Street in London (an entirely dieselized route, restricted to buses and taxis) also demonstrated airway inflammation (myeloperoxidase in induced sputum) and lung function (Forced Expiratory Volume in one second – FEV_1) decrements compared to the responses observed during a low traffic route exposure (Hyde Park). Taken together, these data suggest that the capacity of traffic (specifically diesel) derived PM to elicit oxidative stress at the surface of the lung triggers inflammation and that recurrent cycles of airway irritation/inflammation may underpin many of the respiratory symptoms observed in exposed populations (McCreanor et al. 2007).

1.1.2 Physical and Chemical Composition of Particulate Matter

Airborne particulate matter represents a heterogeneous mixture of solid particles or liquid droplets of varying composition and size. According to US Environmental

Protection Agency, reflecting a general consensus in the literature, ambient PM are classified based on their median aerodynamic diameter as: particles with an aerodynamic diameter less than 0.1µm ($PM_{0.1}$) or ultrafine particles, particles with an aerodynamic diameter <2.5µm ($PM_{2.5}$) also referred to as fine particles and, particles with an aerodynamic diameter between 2.5µm and 10.0µm ($PM_{2.5-10}$), generally classified as coarse particles (Harrison and Yin 2000). In terms of traffic pollution, PM is derived from three separate sources; tailpipe emissions, largely elemental and organic carbon in the ultrafine range; abrasion processes related to mechanical brake and tire wear (in the fine to coarse mode), and the resuspension of road dust in the wake of passing traffic (largely coarse mode) (Thorpe and Harrison, 2008).

Particle size is important as it dictates the deposition pattern within the respiratory tract (**Figure 1.1**). Coarse particles with least surface area, largely deposit in the upper airways by impaction (thoracic PM), whilst fine and ultrafine particles having larger surface area penetrate to the lower airway and deposit in the alveoli (respirable fraction). Although, fine and ultrafine particles penetrate to the distal structures of the lung, the increasing surface area of the lung proximal to distal means that their effective tissue dose is comparatively low relative to particles depositing in the central airways (WHO 2004). Nasal versus oral breathing also affects particle deposition pattern in various regions. The additional deposition of particles in the nose results in an increased extrathoracic (ET) deposition. Consequently, it causes decreased deposition in the bronchial (Bb) and, the alveolar (Al) regions compared to mouth breathing (Kreyling et al. 2004) – **Figure 1.1**.

1.1.3 Air Pollution Abatement Policies

It is now well established that ambient PM impacts on a range of cardiopulmonary health endpoints. The causal link between ambient PM with negative health effects is most strongly illustrated by studies that examined how reductions in pollutant concentrations have resulted in a measurable health dividend. For example, re-examination of the cities initially studied in the seminal Harvard 6-cities study between

1974 and 1982 (Dockery et al. 1993) over the period 1990 to 1998, during which time $PM_{2.5}$ concentrations fell markedly, demonstrated a significant reduction in the risk of pre-mature mortality (Laden et al. 2006). Similarly, a German study covering the period 1991 to 2002 (Breitner et al. 2009) examining the respiratory health of residents living in congested and non-congested neighbourhoods of Erfurt, Germany, found that implementation of air pollution control measures resulted in marked improvements in respiratory health. Similar positive effects on respiratory symptoms following implementation of air pollution control measures have been observed in other studies (Burr 2004; Laden et al. 2006; Renzetti et al. 2009).

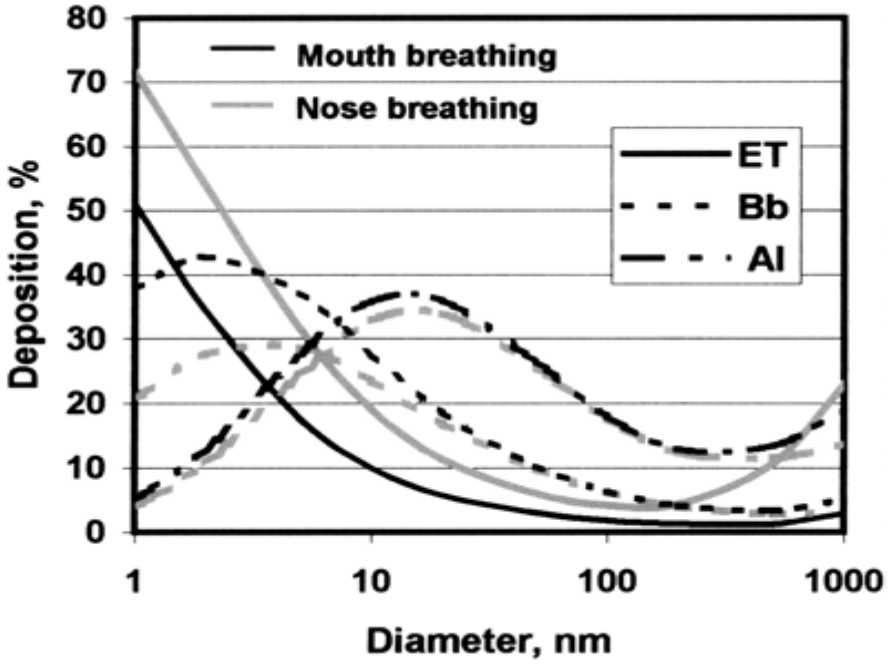


Figure 1.1 Deposition probabilities of inhaled particles according to particle size for nose and mouth breathing in the respiratory tract (extra-thoracic (ET), bronchial airway (Bb) and, the alveolar (Al) compartment). (Source: Kreyling et al. 2004).

These observations, allied to the robust literature linking proximity to traffic and traffic related pollutants to impaired health, have led to an appreciation that regulation of the traffic fleet, both in terms of numbers and emission profile may result in improved public health, especially within the urban environment. Consequently, a

number of countries have introduced Low Emission Zone (LEZ) (**Table 1.1**). These LEZs adopt a variety of measures to regulate traffic emissions, with restrictions on vehicle emissions (related to Euro class), weight, size and type (petrol or diesel) within a designated zone, either by direct restriction or penalty fines for non-compliance. The largest such control zone was introduced into London in January 2008, covering the whole of greater London, 2,644km^2 and a resident population of 8,505,000 (United Nations World Urbanization Prospects, 2005 estimate for London). The impact of the introduction of this zone on the respiratory health of children living in London will represent the major focus of this thesis.

Table 1.1 National Low Emission Zones throughout Europe

Country	No. of LEZ areas	Start Date	Vehicle type restriction	Euro PM size	Retro-type allowed
Austria	1	1/1/2007	Over 7.5T	Euro 2-4	No
Czech Republic	1	1/9/2008	All diesel-based over 3.5T	-	-
Denmark	3, 5*	1/2/2009	Over 3.5T	Euro 3-4	Yes
Germany	32, 10*	1/3/2008	All petrol and diesel-based	Euro 2-4	Yes
Italy	5, 1*	1/1/2007	All vehicles	Euro 1-4	Yes
Netherland	8, 30*	1/7/2007	Over 3.5T	Euro 2-4	Yes
Norway	1*	1/9/2010	Over 3.5T	Euro 4	No
Sweden	5, 4*	1/1/2007	Over 3.5T	Euro 2-3	No
United Kingdom	3, 3*	4/2/2008	Variable	Euro 3-4	Yes

Note: The symbol (*) represents number of LEZ planned or estimated for future.

(Source: modified from Low Emission Zone in Europe website-http://www.lowemissionzones.eu/ overview-of-lezs-othermenu-209).

1.2 Biological Responses to Air Pollution

1.2.1 Induction of Oxidative Stress

Oxidative stress occurs *in vivo* due to imbalance in the prooxidant-antioxidant equilibrium resulting in oxidative damage (Sies 1991). Free radicals are defined as any species containing at least one unpaired electron capable of existing as individual entity (Halliwell 1991) and include the hydroxyl radical (OH•), the superoxide anion (O_2•$^-$),

and nitric oxide (NO) (**Figure 1.2**). Reactive Oxygen Species (ROS) include free radicals, but also contains reactive chemical species, without unpaired electrons, which are capable of participating in redox reactions, such as hydrogen peroxide (H_2O_2). The production of these species can result in damage to proteins, lipids and nucleic acids resulting in the disruption of normal cell function and ultimately cell death (Kidd 1997). It should be noted however that at low concentrations many ROS actually appear to perform regulatory roles in cells and that in certain cases even excess production of these species may be beneficial, e.g. oxidant production by phagocytic inflammatory cells during the respiratory burst.

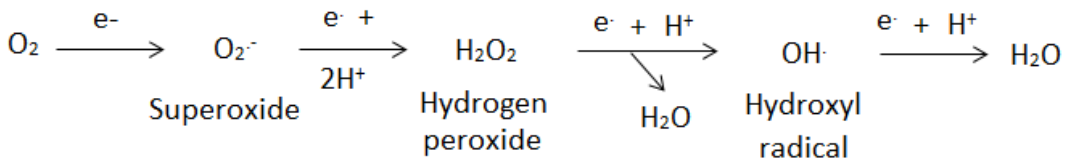


Figure 1.2 Pathway of oxygen reduction, illustrating the formation of the major physiologic reactive oxygen species. (Adapted from Li et al. 2003).

1.2.1.1 Sources of Reactive Oxygen Species

PM-induced ROS generation may arise through a variety of pathways: the introduction of redox catalysts (Fe, Cu, quinones) into the lung, the xenobiotic metabolism of polyaromatic hydrocarbons (PAH), the induction of active inflammation and the disruption of the mitochondrial electron transport chain (Ayres et al. 2008). These will be discussed in the following sections.

1.2.1.1.1 Introduction of Redox Catalysts to the Lung

Redox-active transition metals are present on the surfaces of ambient PM and include iron (Fe), copper (Cu), vanadium (V), nickel (Ni) and chromium (Cr). Their significance lies in the fact that they will catalyze the production of free radicals and

$O_2^{-\cdot} + Fe^{3+} \longrightarrow Fe^{2+} + O_2$ Haber - Weiss Reaction (Equation I)

$H_2O_2 + Fe^{2+} \longrightarrow Fe^{3+} + OH^- + \cdot OH$ Fenton Reaction (Equation 2)

ROS *in vivo* (Comhair and Erzurum 2002) through Fenton and Haber–Weiss like reactions (Equation 1).

Similarly, PM- associated quinone compounds produced during fuel combustion are able to redox cycle *in vivo* with the production of superoxide, with the degree of cycling dependent on the availability of NADPH-CYP450 reductase (Sioutas et al. 2003). Notably quinone-induced oxidative stress can also occur as a consequence of the metabolism of PAHs associated with inhaled PM. PAHs are metabolized *in vivo* by three pathways involving action of CYP450 1A1, expoxide hydrolase, and dihydrodiol dehydrogenase, with CYP450 1A1 expression under the regulation of the aryl hydrocarbon receptor (AhR). The initial oxidation of particle-associated PAHs is catalyzed by CYP450 1A1 producing PAH dihydrodiol, which is then either hydrolysed by expoxide hydrolase producing PAH diol-epoxide, or acted upon by dihydrodiol dehydrogenase to produce catechol, which is subsequently converted into quinone. The quinone by the action of NADPH-P450 reductase and removal of one electron is converted into semiquinone, which can be converted back to quinone initiating redox cycling. In this process the electron is donated to oxygen producing $O_2\bullet^-$. This process has been implemented as the basis for the generation of ROS in cells challenged with diesel exhaust particles, which upon inhalation introduce PAHs into the lung, as illustrated in **Figure 1.3**.

1.2.1.1.2 ROS Production during Inflammation

When the lung is challenged by foreign material such as inhaled PM, the body's defense system responds by initiating an acute inflammatory response, usually characterized by the migration and functional activation of phagocytes and lymphocytes. Under normal circumstances, this is a beneficial response, but aberrant over activation of this response, such as occurs during PM-challenge can result in significant tissue injury. Consistent with this scheme, healthy humans exposed to near ambient concentrations of diesel-derived PM_{10} have been shown to elicit an acute airway inflammation, with up-regulation of vascular adhesion molecules, production of

pro-inflammatory cytokines and increased airway neutrophils, mast cell and T cell numbers (Salvi et al. 1999). This acute inflammatory response examined 6 hour post-challenge, has been associated with the activation and nuclear migration of redox-sensitive transcription factors such as nuclear factor KB (NF-KB) and activator protein 1 (AP-1) (Pourazar et al. 2009). As these transcription factors drive the expression of a panel of pro-inflammatory mediators, these data have provided evidence for PM-induced oxidative stress being the trigger to the induction of airway inflammation.

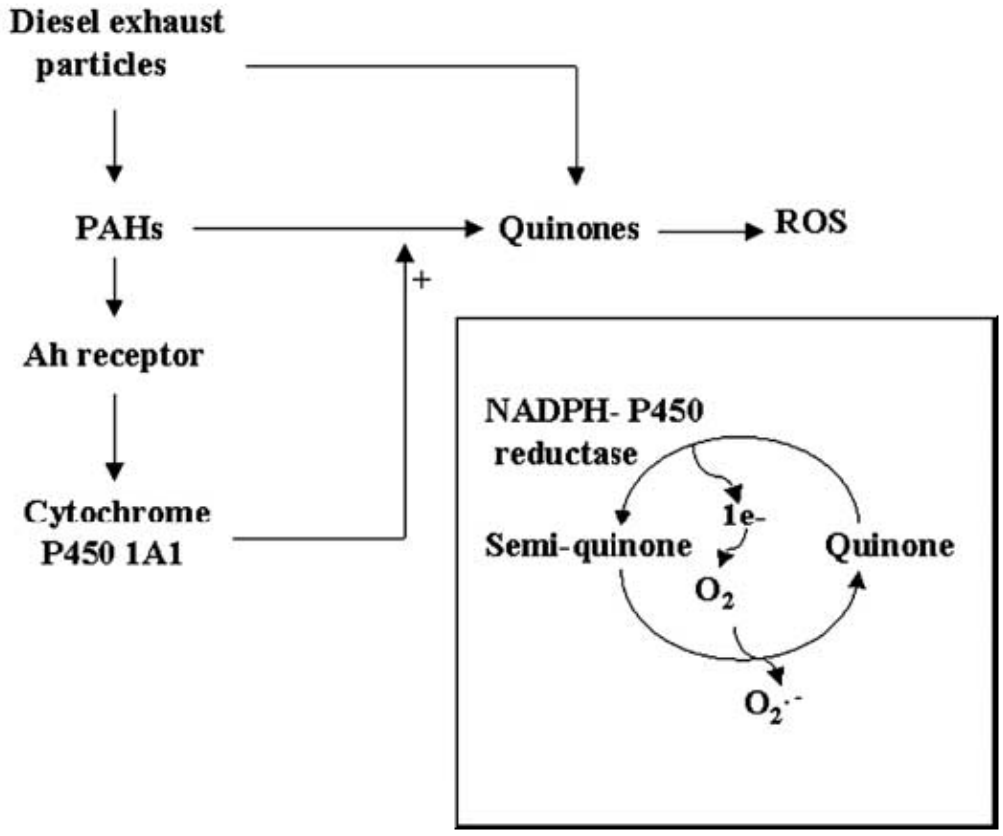


Figure 1.3 Quinone formations via PAH metabolism and the subsequent production of ROS. (Source: Li et al. 2003).

1.2.1.2 Pulmonary Antioxidant Defense Mechanisms

The lung possesses a robust antioxidant defense network to protect itself from oxidative injury. These include both low molecular weight (glutathione, α-tocopherol, ascorbate and urate) and enzymatic antioxidants (superoxide dismutases (SOD), catalase, peroxidases, peroxyredoxins, thioredoxins, and the glutaredoxins) distributed between the cellular and extra-cellular lung compartments. The first line of defense against inhaled materials resides within the respiratory tract lining fluids (RTLFs) that bathes the respiratory epithelium throughout the airways (Rahman et al. 2006; Wright et al. 1994).

1.2.1.2.1 Low Molecular Weight Antioxidants

This group act directly on the pro-oxidants by donating an electron (hydrogen) thus elminating ROS. In this process, they are oxidized to low reactivity free radicals themselves. Glutathione (GSH) the major cellular low molecular weight antioxidant functions in the following ways: (i) by acting as a substrate for glutathione peroxidase, (ii) as a direct scavenger of ROS and (iii) by permitting the reduction of dehydroascorbate back to ascorbate by acting as a substrate for dehydrascorbate reductase (Townsend et al. 2003)**.** Ascorbate itself can also act as a scavenger by donating an electron to ROS, forming the ascorbyl radical, which can undergo further oxidation to dehydroascorbic acid. However, ascorbate can also react with ferric ions to produce the highly reactive OH• radical. Thus, whether ascorbate functions as a pro- or anti-oxidant largely depends on its concentration relative to the availability of free Fe or Cu in the cell or extra-cellular fluid (Romieu et al. 2008). Urate produced by the oxidation of xanthine and catalysed by xanthine oxidoreductase has also been shown to be an important sacrificial antioxidant *in vivo* especially against reactive nitrogen species and the hydroxyl radical. There has also been speculation that urate may chelate iron, thus limiting deleterious Fenton chemistry, but this has yet to be confirmed *in vivo* (Kohen and Nyska 2002).

1.2.1.2.2 Enzymatic Antioxidants

Three types of superoxide dismutase exist in humans: extracellular-SOD (EC-SOD), intracellular Cu/Zn-SOD and Mn-SOD. These enzymes function to remove superoxide by catalyzing its dismutation to H_2O_2. The H_2O_2 thus formed can then be further reduced to water by the action of catalase and glutathione peroxidase (GPx). The extra-cellular form of this enzyme is expressed by alveolar macrophages, bronchial epithelium, extracellular matrix and epithelial cells and appears particularly important in protecting the lung from oxidative injury. Due to its hydrophobic nature, EC-SOD binds to lipophilic cell surfaces by binding to type I and type IV collagen through its matrix binding sites (Bowler et al. 2002). In contrast, the expression of Mn-SOD is restricted to the mitochondrial matrix where it functions to limit injury from superoxide formed as electrons leak from proteins within the electron transport chain, whilst Cu/Zn-SOD has a more widespread cellular distribution.

Catalase and GPx are involved in the removal of biological peroxides, which otherwise would become substrates for the production of more damaging ROS. Catalase located within cellular peroxisomes in macrophages, pneumocytes and fibroblasts catalyses the conversion of H_2O_2 to water. Low expression of catalase and SOD has been associated with increased levels of oxidative stress (Matés 2000). Four selenium-based cellular and extracellular GPx have been identified and are named according to their cellular location: (i) cytosolic, (ii) membrane-associated GPx, (iii) cytoplasmic GPx and, (iv) gastrointestinal GPx. These enzymes catalyse the removal of H_2O_2 and large molecular lipid peroxides. The reducing action of GPx requires GSH, which is converted to glutathione disulphide (GSSG) during this reaction. The formed GSSG can subsequently be converted back to GSH by glutathione reductase using NADPH as an electron donor (Lu 1999).

1.2.2 Induction of Xenobiotic/Antioxidant Responses by Inhaled Pollutants

The initial response of cells/tissues exposed to oxidative stress via the inhalation of airborne xenobiotics is the induction of antioxidant and xenobiotic metabolizing enzymes. This occurs through the transcriptional activation of the antioxidant responsive element (ARE) by the transcription factors nuclear regulatory factor 2 (Nrf2) and activator protein-1 (AP-1). Nrf2/AP1 induction of ARE genes (heme oxidase-1 (HO-1), NADPH quinone-oxidoreductase (NQO1), cytochrome P450 1 alpha 1 (CYP450 1A1) (phase I), glutathione s transferase (GST)) represents the initial response to oxidative stress, and it has been suggested that variations in the capacity to mount these responses – protective adaptations – may play a role in individual susceptibility to air pollutants (Cho & Kleeberger 2010). If these initial adaptation reactions are insufficient to 'cope' with the oxidative insult the response shifts to a more pro-inflammatory profile, with activation of mitogen-activated protein (MAP) kinase cascades and NFκB, driving cell proliferation and the expression of pro-inflammatory mediators (**Figure 1.4**). Finally, if the oxidative stress is sufficiently severe, cell arrest occurs (to allow DNA repair) and ultimately cell death (under the regulation of p53) will be triggered via apoptosis or necrosis (Gilmour et al. 2006). This hierarchical response model is illustrated in **Figure 1.4**, with the key regulatory pathways highlighted, corresponding to cellular adaption, inflammation and cell death. The scheme outlined overleaf has also been modified to illustrate the 'initial' tolerance of the airway to inhaled oxidants due to the endogenous aqueous antioxidants within the respiratory tract lining fluid.

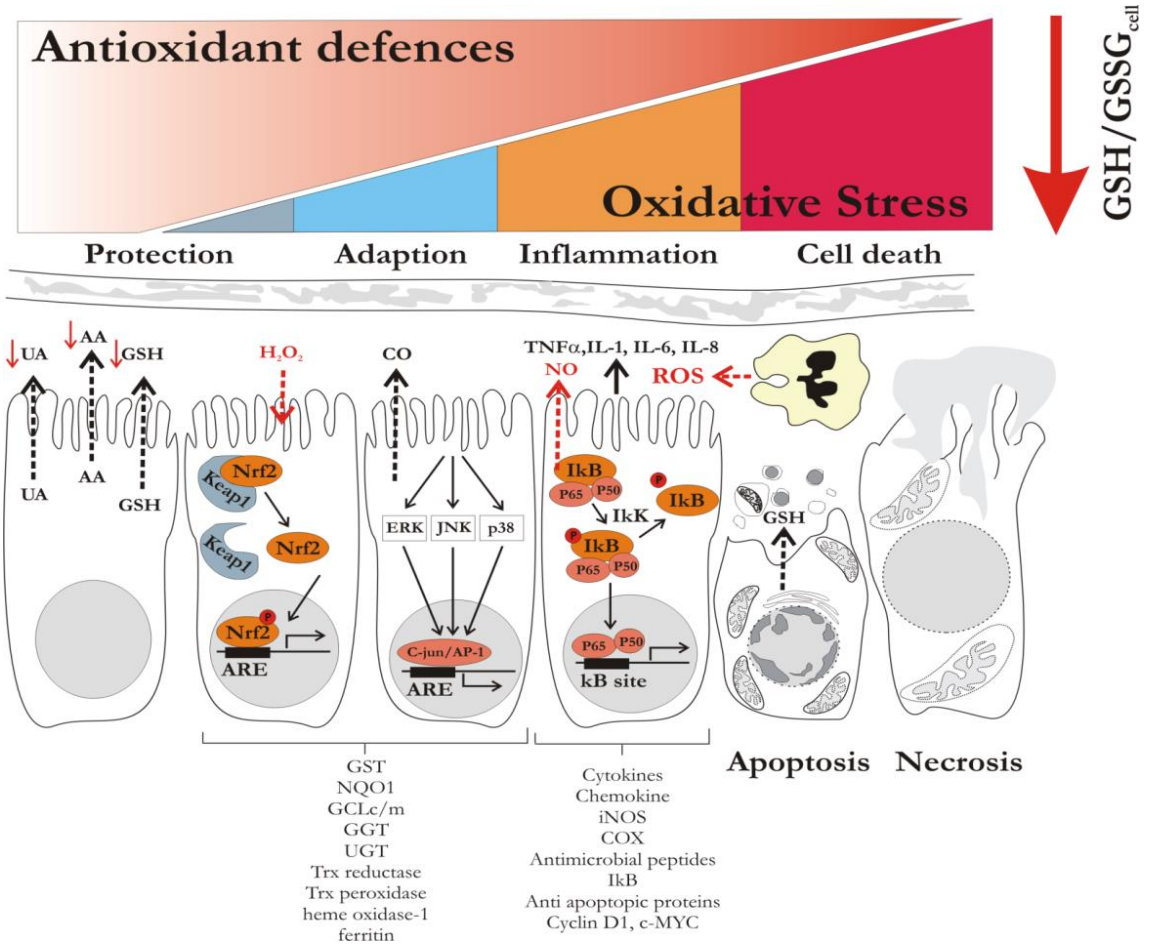


Figure 1.4 Diagrammatic representation of the hierarchical response of cells to PM-induced oxidative stress at the air-lung interface**.**

This diagram represents an extension of the earlier model proposed by (Xiao et al. 2003) to include the influence of the antioxidant defenses within the RTLF. In this modified model, the initial defense against PM-induced oxidative stress resides within the RTLFs, initially characterized by acute early losses of ascorbate (AA), urate (UA) and glutathione (GSH). In light of the previous findings, we suggest initial losses due to the inhalation of pollutant gases and particulates are partially offset by movements of these antioxidants, particularly GSH from the epithelium to the extra-cellular compartment (Mudway et al. 2006; Behndig et al. 2009; Sehlstedt et al. 2010). When these defenses are overwhelmed, the underlying cells initially induce adaptive strategies, under the regulation of Nrf2 and AP-1 to deal with the oxidative stress, or process the inhaled xenobiotics. The figure illustrates a number of genes known to be up-regulated in these adaptive responses, under the regulation of Nrf2 and AP-1, including the NQO1, SOD, HO-1, and GCL. Further oxidative stress, reflected by a decrease in the cellular GSH/GSSG ratio leads to the transcription of genes under the regulation of NFκB, such as pro-inflammatory cytokines (TNF-α, IL-6, and IL-8) and inducible nitric oxide. Further oxidative stress subsequently leads to cell arrest and induction of cell death, either by apoptosis or by necrosis.

1.2.2.1 Cytochrome P450

Cytochrome P450s are heme-containing enzymes, which catalyze Phase I metabolism reactions. Cytochrome P450 CYP1A1 is one of the three members of the CYP1 family, located mainly in extrahepatic tissues, which participates in the detoxification of numerous xenobiotics, including an important role in PAH metabolism. The metabolism of the PAHs is initiated by their binding to the cystolic aryl hydrocarbon receptor, which then translocates to the nucleus, where it heterodimerizes with the aryl hydrocarbon nuclear translocator protein forming a complex which interacts with AhREs (Aryl hydrocarbon response elements), XREs (Xenobiotic response elements) or DREs (Dioxin response elements), to activate gene transcription (Hankinson 1995). The human CYP1A1 gene contains many xenobiotic responsive elements (XRE) that suppport AhR mediated transcription. CYP1A1 transforms PAHs to highly reactive diol epoxides, which are subsequently detoxified by phase II enzymes, including epoxide hydrolases and glutathione S-transferases (e.g. GSTM1 and GSTP1). The potency of diesel exhaust particles and their organic extracts to induce gene expression of CYP1A1 has previously been demonstrated in human lung samples ex vivo (Iba et al. 2010) and in human airway epithelial (16HBE) and human macrophage (U937) cell lines (Bonvallot et al. 2001, 2002; Vogel et al. 2005).

1.2.2.2 Glutathione S Transferases (GSTs)

GSTs are important phase II isoenzymes, expressed both in the cystolic and membrane compartments of the cell. The cytosolic enzymes are dimeric and soluble, whilst the microsomal protein enzymes are trimeric in structure. Although, humans contain 16 such soluble enzymes, arranged in 8 groups, the major focus in the air pollution literature has been on the mu (GSTM1), pi (GSTP1) and theta (GSTT1) types. The main function of the GSTs is the detoxification of electrophiles by conjugation to glutathione. This includes the conjugation of substrates such as oxidized PAH epoxides, derived from action of phase 1 cytochrome P-450s. GSTP1 catalyses the conjugation of the diol epoxide of benzo(a)pyerene, benzo(c)phenanthrene and benzo(g)chrysene. In

contrast, GSTT1 can activate dichloromethane (DCM) by conjugation with GSH to form S-chloromethylglutathione. In addition, GSTM1 and GSTP1 cause detoxification of comparatively bulky aryl halides, whereas GSTT1 catalyzes small dihaloalkanes (Hayes and Strange 2000).

1.2.2.3 NADPH-quinone Oxidoreductase (NQO1)

NQO1 is a phase II antioxidant enzyme, expressed in the cytosol of non-phagocytic cells, polymorphonuclear leukocytes, monocytes and macrophages. It provides protection against quinone induced oxidative stress by reducing them to redox-stable hydroquinones, in a single two-electron step utilizing either NADH or NADPH as a reducing cofactor. The redox stable hydroquinones can be quickly conjugated and excreted (Ross et al. 2000). NQO1 can also catalyze the bioreductive alkylation of quinones producing products capable of alkylating nucleophilic sites (Reddy 2008). NQO1 can catalyze reactions involving substrates such as ortho and para-quinones, menadione, quinone-imines, benzoquinone, dichlorophenolindolphenol and naphthoquinone. The reaction involves a hydride transfer between the NADH and flavin adenine nucleotide (FAD) cofactors and from $FADH_2$ to the quinone substrate. Menadione conversion to hydroquinone by NQO1 using two-electron reduction has a direct competition with the cellular reductases using one-electron reduction (Hoidal 2001).

1.3 Population Susceptibility to Air Pollution Induced Health Effects

Assessing an individual's susceptibility to an inhaled xenobiotic depends on multiple intrinsic (internal or host) and extrinsic (external) factors and their relationship to the genotype (Kleeberger 2005). The intrinsic or host factors include gender, age, diet, and ethnicity, whilst the external factors reflect social factors including provision of health care. The healthcare related factors include health behaviors and differential access to health resources (Lurie and Dubowitz 2007). In contrast, social factors include

socioeconomic status, exposure to environmental triggers such as allergens, air pollutants and tobacco smoke, family history, lifestyle, and culture (Collins 2004; Drake et al. 2008).

1.3.1 Susceptibility of Children

Generally, children appear more susceptible to respiratory ailments than adults upon exposure to environmental pollutants. This disparity arises from physiological, metabolic and behavioral differences between the two populations (Bateson and Schwartz 2008; Cohen et al. 2000; Kimmel et al. 2005; Landrigan et al. 2003; Rogan 1995). Children are born with undeveloped lungs, characterized by an immature epithelium and significantly fewer alveoli (20-50 million) than adults (Burri 1984; Dunnill 1962; Langston et al. 1984). However, a significant increase in the number of alveoli occurs during first year, reaching 300 million at the age of eight years (Burri 2006; Reid 1977). During lung morphogenesis, the multiplication of alveoli begins, in addition to the branching of the respiratory bronchioles and the differentiation of epithelium (Joshi and Kotecha 2007; Kotecha 2000). Therefore, any air pollutant-related damage occurring to the developing lungs at this stage are likely to result in airway remodeling leading to changes in morphology persisting into adulthood. Animal studies have shown altered epithelial development and differentiation after just a single environmental insult during early life. In one such study exposure of neonatal mice to naphthalene resulted in enhanced epithelial injury compared with adult mice (Fanucchi et al. 1997, 2004).

Additionally, children have a higher basal breathing rate and their lungs have a larger surface area per unit body weight compared to adults (Gagliardi and Rusconi 1997). Hence, the volume air/unit body mass passing through the lungs of an infant is approximately double than that of an adult, which will clearly result in an enhanced uptake of inhalable pollutants (Arcus-Arth and Blaisdell 2007; Iliff and Lee 1952; Poets et al. 1993). Children's breathing rate decreases steadily during growth, implying that newborns would be most susceptible to inhaled xenobiotics, followed by infants, with

susceptibility decreasing with age (Plopper et al. 1994; Fanucchi et al. 1997; 2004; Gehring et al. 2002; Pierse et al. 2006). Children and adults also show differences in patterns of breathing which may affect the deposition of inhaled pollutants. At rest, children normally breathe through the mouth. In contrast, adults are obligate nasal breathers (Bennett and Zeman 2004). As the nose effectively filters larger particles from the air stream (Nikasinovic et al. 2003; Bennett et al. 2003; 2008), one might consider mouth-breathing to further enhance pollutant uptake into the airways. Consequently, low pollutant exposures that are largely asymptomatic in adults, may cause airway irritation in children.

In addition to the biological basis for increased susceptibility of children to air pollutants, it should also be noted that their exposures are elevated compared with adults due to the increased period of time they spend exercising in the outdoor environment. Children while playing, experience rapid and deep breathing, which not only increases the inhaled dose of any air borne pollutants but also increases deposition to the distal lung. This increased distal dose is compounded by the slower rate of particle clearance from region (Ekelund et al. 2004; Ginsberg et al. 2005).

From the evidence outlined above, there is a sound basis for believing that children are more susceptible than adults to air pollution hazards. This contention is supported by a large number of air pollution studies that have demonstrated strong association between exposures to ambient air pollutants with substantial increased rates of respiratory symptoms such as chronic cough, bronchitis, substantial decreased of lung function, airway sensitization, and increased incidences of respiratory infections in children (Brunekreef et al. 1997; Venn 2000; Brauer 2002; Gehring et al. 2002). More recent cohort based studies have also demonstrated associations between long term traffic related pollutant exposures with impaired respiratory health in children (Gauderman et al. 2007b; Lindgren et al. 2009; Migliore et al. 2009; Morgenstern et al. 2008; Nordling et al. 2008; Oftedal et al. 2009; Rosenlund et al. 2009; Schultz et al. 2012). In many of these studies, the proximity of the schools the children attend to busy roads have been employed as a surrogate for pollutant exposure, and were related to a range of respiratory and allergic symptoms (Brauer et al. 2002; Gauderman et al. 2005;

Nordling et al. 2008). In addition, traffic density and type has also employed as an exposure surrogate in certain studies (Janssen et al. 2003; Ryan et al. 2005). A smaller panel of studies have actually employed direct pollutant measurements at the schools (Brunekreef et al. 1997), or the children's residence (Duhme et al. 1996; Gehring et al. 2002; Nordling et al. 2008), with modeled pollutant concentrations employed in larger cohort studies (Brauer et al. 2007; Morgenstern et al. 2008; Ryan et al. 2007).

1.3.1.1 Respiratory Airway Disease

Numerous studies have demonstrated clear and robust associations between a worsening of cardiopulmonary symptoms in adults and children with increased exposures to traffic related air pollutants. Whilst the non-allergic respiratory symptoms such as mortality (Bobak and Leon 1999; Glinianaia et al. 2004) and birth weight (Bobak et al. 2001) reported in relation to traffic related air pollutants appear largely consistent amongst the studies performed to date. However, there is inconsistency in the results of studies examining the relationship between traffic pollutants and allergic symptoms: asthma, wheeze, allergic rhinitis and sensitization. This will be discussed in greater detail in the sections below.

1.3.1.1.1 Asthma and Wheeze

Whilst numerous studies illustrate positive associations between exposures to traffic derived pollutants with a worsening of asthma symptoms (Brauer et al. 2002; Duhme et al. 1996; Gauderman et al. 2005; Morgenstern et al. 2007, 2008) other studies have failed to confirm this observation (Ciccone et al. 1998; Migliore et al. 2009; Nordling et al. 2008; Oftedal et al. 2009). The reason for the lack of association observed in some studies is unclear, though it may reflect inconsistency in the methodologies employed between the different studies. Most of the studies diagnosed asthma based on self-reported wheeze, rather than an accurate doctor's diagnosis, or evidence of airway reversibility upon salbutamol treatment. In addition, most of the

studies focused on young children where the diagnosis of asthma is challenging. Similarly to the asthma associations reported above, the association between wheeze and traffic pollution was inconsistent, with many studies demonstrating significant associations (Brauer et al. 2007; Brunekreef et al. 1997; Duhme et al. 1996; Nordling et al. 2008; Ryan et al. 2005), whilst others were unable to confirm this observation (Ciccone et al. 1998; Venn et al. 2000; Wyler et al. 2000).

One statistically powerful way of investigating the relationship between exposures to traffic related pollutants and respiratory health has been to examine indices of respiratory morbidity in established birth cohorts and to relate these observations to modeled or measured pollutant concentrations over time. One such Dutch and German birth cohort TRAPCA (Traffic Related Air Pollution on Childhood Asthma) with 4,000 children, aged 2 years and 1,756, aged 1-2 years has shown significant associations between traffic pollutants and the incidence of respiratory diseases (Brauer et al. 2002; Gehring et al. 2002). To improve the linkage between respiratory symptoms and pollutant exposures, Brauer et al. (2007) extended the initial analysis on this birth cohort to assess individual annual exposure to NO_2, $PM_{2.5}$, and soot using geographic information system (GIS) related to the cumulative incidence of reported symptoms at the age of 4 years. Elevated odd ratio for wheeze (1.2 (1.0-1.4)) and doctor diagnosed asthma (1.3 (1.0-1.7)) were observed in children per inter-quartile range of modelled soot exposure ($0.58\times10^{-5}\,m^{-1}$). To further refine the likely actual exposure of children to traffic-related pollutants and to relate this to respiratory health, Morgenstern et al. (2007) performed a study including all the children (1-2 years) in the entire Munich metropolitan area, which employed buffer zone variables such as distance from the main road (50m) to further improve the individual exposure assessment. The results from this study refined the initial observations with stronger associations between respiratory symptoms and traffic related pollutants exposures noted in girls compared with boys (Morgenstern et al. 2007).

Further research has demonstrated a correlation between increased freeway traffic, especially diesel powered heavy goods vehicles with asthma prevalence in children (Duhme et al. 1996; Brunekreef et al. 1997; Gauderman et al. 2005).

Subsequent studies, specifically the Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS) have demonstrated that traffic density/volume appears to be less of a risk for poor respiratory health (wheezing) than the distance from the traffic source, the type of traffic, or driving scenario: major highways, stop and go, HGVs, and bus traffic (Ryan et al. 2005). Infants exposed to stop and go bus and truck traffic demonstrated a significant increased risk of wheeze compared to unexposed infants. However, infants living nearby high volume traffic did not show significantly increased rates of wheezing. In their later study, Ryan et al. (2007) improved the exposure assessment by using land use regression (LUR) and reported a significant increased risk of wheeze associated with elemental carbon attributable to traffic sources within range of 0.30-0.90 ug/m^3 (Ryan et al. 2007).

Gauderman and colleagues (2005) assessed the effect of multiple indicators of traffic/pollutant exposure on the doctor-diagnosed asthma and wheezing in 208 children from 10 southern California communities. The indicators chosen included NO_2 exposure, season and residential distance to the nearest freeway. The lifetime freeway exposure and measured NO_2 concentrations were found in this study to be associated with doctor-diagnosed asthma and wheezing, however, no association was found with traffic volume in accord with the findings of Ryan et al. (2007).

The Swedish BAMSE (Barn, Allergy, Milieu, Stockholm, Epidemiological) study, conducted in Stockholm and involving approximately 3,500 children followed from birth to 4 years of age also examined the relationship between respiratory health and allergic sensitization with exposure to traffic (Nordling et al. 2008). Traffic-NO_x exposures, based on modeled concentrations were shown to be significantly associated with persistent wheeze (wheeze both at one and four years of age) only in females, but not with doctor-diagnosed asthma. In addition, the association between NO_x and wheeze was stronger in non-atopic children. Similar but insignificant associations were demonstrated for PM_{10}.

1.3.1.1.2 Allergic Sensitization

Traffic-related allergic sensitization, assessed either by skin prick or blood testing in the urban areas has been reported in a number of studies (Ciccone et al. 1998; Nordling et al. 2008), however this association is far from robust (Nicolai et al. 2003; Oftedal et al. 2007; Rosenlund et al. 2009). Studies have reported that children living in proximity to major roads exhibit impaired respiratory health, especially sensitization to pollen (Kramer et al. 2000; Wyler et al. 2000; Janssen et al. 2003; Morgenstren et al. 2008). Janssen et al. (2003) investigated the effect of traffic density and composition on the allergic symptoms of children (2,500, aged 7-12 years) attending schools near roads (407±230 m) in The Netherlands. In this study, a greater risk of sensitization to pollen was observed in relation to exposure to diesel powered heavy goods vehicles, but not car traffic counts. This association was further supported by the results of a birth cohort study performed in the Munich metropolitan area examining allergic sensitization to pollen during the first six-years of life in 3000 children (Morgenstern et al. 2008). In this study, exposure to elevated $PM_{2.5}$ and residential proximity to busy roads (less than 50m) was associated with significant increased risk of sensitization to pollen (outdoor) allergens. In addition, the distance from the main road had a dose-response relationship i.e. the greater the distance from a road, the lower the $PM_{2.5}$ exposure and the lower the degree of sensitization observed. Furthermore, increased production of IgE antibodies has been associated with elevated exposures to PM_{10} and NO_x in children at 4 years of age (Nordling et al. 2008). Overall, these results indicate that children living near busy roads are more likely to become sensitized by pollen. Possibly, due to pollen and air pollutant interactions.

1.3.1.1.3 Allergic Rhinitis (hay fever)

Several studies have also examined the relationship between traffic pollutant exposures and the incidence of hay fever. Again, as with the wheeze and asthma data whilst many studies have demonstrated significant interactions (Duhme et al. 1996; Lee

et al. 2003; Morgenstern et al. 2007, 2008) others have not (Ciccone et al. 1998; Krämer et al. 2000; Nicolai 2002) despite the application of standard procedures in most studies. Kramer et al. (2000) reported associations between hay fever with traffic-related, outdoor NO_2, but not indoor sources of this gas. Lee et al. (2003) conducted a similar but larger study involving approximately 330,000 school children in Taiwan, with the actual application of monitored CO, NO, SO_2 and PM_{10} from the pollution stations located within 2 km of the schools. They observed significant associations between doctor-diagnosed allergic rhinitis and increased concentrations of CO, NO, SO_2 and PM_{10}. The extended study of TRAPCA II (Morgenstren et al. 2007) also showed significant associations of $PM_{2.5}$ with sneezing, runny/stuffed nose during the first two years of life, consistent with the previous research (Gehring et al. 2002).

1.3.1.1.4 Atopic Dermatitis (eczema)

Studies have also been performed examining the relationship between traffic exposures and the presence of dermatitis in children, with mixed results. Whilst several studies have reported a significant association between the presence of childhood eczema with traffic exposures (Ising et al. 2003; Morgenstern et al. 2008) others have failed to confirm this linkage (Brauer et al. 2002, 2007; Kramer et al. 2000).

1.3.1.1.5 Lung Function

Several cohort studies have assessed the effect of traffic pollutants on measures of lung function, including forced vital capacity (FVC), forced expiratory volume (FEV) and peak expiratory flow (PEF) in children (Barraza-Villarreal et al. 2008; Brunekreef et al. 1997; Dockery et al. 1996; Gauderman et al. 2000, 2007b; Nordling et al. 2008; Oftedal et al. 2008; Rosenlund et al. 2009). A more detailed account of the more recent literature will be provided in **Chapter 3**. A study in Rome (Rosenlund et al. 2009) investigated respiratory symptoms of 2,107 children

(aged 9-14 years) from 40 schools, with spirometry performed on 1,359 children. Three indicators were investigated which included self-reported heavy traffic outside home, measured distance between home and busy roads and the residential NO_2 concentration in the area. The results indicated significant association between the NO_2 concentration and reduced expiratory flow, FEF_{25-75}, FVC and PEF. The impact of ambient pollution ($PM_{2.5}$) on the lung function of children was also assessed in 12 communities in Southern California over a period of 4 years (Gauderman et al. 2000) with evidence of a reduced FEV_1 in children from polluted areas consistent with retarded lung growth. Additionally, in a later study, Gauderman et al. (2007) assessed the effect of the local exposure of major roads on lung function, demonstrating that children living within 500m of freeways have a substantial reduction in lung growth (FEV_1, –81 mL; MMEF, -127ml) compared with children living at least 1,500m away from the pollutant source. A significant association between traffic NOx and PM_{10} with the reduced lung function during the first year of life has also been reported (Nordling et al. 2008). In these studies, long term pollutant averages were used to approximate difference in life time pollutant exposures which is clearly an imperfect indicator of personal exposure. Therefore, Oftedal et al. (2008) calculated hourly concentrations of NO_2, $PM_{2.5}$ and PM_{10} to access individual exposure in children living in Oslo, Norway, since birth. Early and lifetime exposures to air pollution were associated with reduced lung growth especially in girls. In this analysis, NO_2 showed both short and long term effects on lung function, whilst $PM_{2.5}$ and PM_{10} was only associated with long-term effects.

1.3.2 Genetic Susceptibility

There is an increasing body of evidence demonstrating that a considerable proportion of the variation in the sensitivity of individuals to air pollutant insults can be attributed to polymorphisms in key genes related to xenobiotic and antioxidant defense, as well as the induction of inflammation. Single nucleotide polymorphisms (SNP) may occur in one of the following areas of the genome *i.e.* coding regions (cSNPs), regulatory regions (rSNPs) or 'junk' DNA regions and are therefore referred to as

anonymous SNPs. **Table 1.2** shows the definition of certain important terms related to genetic polymorphisms.

Genetic contribution to the complex diseases is still not fully understood. However, it is believed that genetic contribution to disease is multi-factorial and may involve many common alleles, which differ in allelic frequency amongst major human sub-populations. Therefore, genetic variance has become an important intrinsic disease determinant and recent research has focused on the identification of genes involved in the numerous health disorders. However, due to the presence of polygenetic factors, the actual disease aetiology is yet to be established, but gene–gene and gene-environment interactions are believed to be the main causes.

Table 1.2 Terms employed to describe genetic polymorphisms

Term	Definition
Genotype	The genetic constitution (the genome) of a cell, an individual or an organism, which is distinct from its expressed features, or phenotype.
Allele	One of the different forms of a gene that can exist at a single locus. Often implies a detectable phenotype difference, which may arise from natural or experimentally induced variation.
Polymorphism	Genetic variation at a locus present at a frequency of at least 1% in a population. Naturally, occurring variations in the DNA sequence and may not have detectable phenotypic variations.
Single nucleotide polymorphism (SNP)	Genetic variation that arises from a change in a single nucleotide at a locus and can be used as a marker for association mapping to assess genetic variation within and among populations.
Coding SNPs (cSNPs)	Present in the actual gene-coding region of a chromosome and have a higher probability of influencing propensity to disease than SNPs found outside gene regions.
Regulatory SNPs (rSNPs)	These SNPs affect regulatory regions that govern gene expression. Thought to be relatively uncommon and potentially as valuable as cSNPs.
Protein SNPs (pSNPs)	When a cSNP or a rSNP leads to an altered amino acid, which in turn leads to altered protein function or expression and an observable change in the organism's phenotype.
Synonymous SNP	A change in a coding region that alters the nucleotide sequence of a codon without changing the amino acid.
Non- synonymous SNP	A change in a coding region that alters the nucleotide sequence of a codon which cause changing in amino acid. This could be missense or nonsense mutation.
Anonymous SNP	No known effect on gene function. Thought to be the most common type of SNPs and possibly valuable as markers for linkage disequilibrium studies, when they are relatively close to the gene

1.3.2.1 Systematic Review on Genetic Association Studies

Systematic literature search was conducted for the identification of effect modification of antioxidant and xenobiotic genes on impact of air pollution on respiratory diseases or symptoms until year 2012. This exercise was undertaken to identify genetic variations, which have been associated to traffic derived pollutants and respiratory disease regardless of populations age. The search engines like PubMed and Scopus were utilized to search information related to air pollution using terms: "air pollution", "nitrogen dioxide", "PM", "ozone"; for genes using terms: "polymorphism", "genetic", "SNP", "antioxidant", "xenobiotic", "GSTP1", "GSTM1", "NQO1", "Nrf2", "AhR", "CYP1A1", "EPHX1"(microsomal epoxide hydrolase 1), "GCLM" and "GCLC" (glutamate-cysteine ligase modifier subunit and catalytic subunit), and for respiratory symptoms using terms: "lung function", "asthma", "respiratory" and "wheezing". **Tables 1.3** and **1.4** show summary of the published studies examining specific genetic variations in relation to effect of air pollutants and respiratory responses for adults (14 studies) and children (19 studies) respectively in chronological order. Some studies investigated the effects of SNPs conjointly with genetic variants other than xenobiotic and antioxidant genes (Li et al. 2006; Melén et al. 2008) such as TNF-α. In addition, due to lack of studies for association between gene of interest such as CYP1A1 and respiratory disease, the studies were still included to understand the mechanism involved in relation to air pollution and the genetic variance. There may be some studies, which are not included mainly due to repetition from the same group. Several studies have been performed using the Children's Health Study population (Breton et al. 2011; Islam et al. 2009, 2008; Li et al. 2006; Salam et al. 2007). The last major study examining the interaction between pollutants and genotype in adults was published in 2010 whereas more genetic association studies have been performed in children since this period and will be reviewed in more detail in the results chapters. The increased interest in childrens research may reflect the established vulnerability of children exposed to ambient pollution. Moreover, studies have predominantly focused on ozone exposure in both populations, rather than primary traffic derived pollutants.

Overall, the GST supergene family which comprises of mu, pi and theta (GSTM1, GSTP1 and, GSTT1) has been investigated the most extensively. These phase II xenobiotic metabolizing enzymes are critical in detoxifying the carcinogenic and ROS producing electrophiles in order to protect lung against insult and prevent further oxidative stress. The GSTM1 genotype, which is located on chromosome 1p13 appears to be the most intensively studied with twelve studies in children (Buthbumrung et al. 2008; David et al. 2003; Hong et al. 2007; Islam et al. 2009; Lee et al. 2008; Li et al. 2006; Piacentini et al. 2010; Reddy et al. 2012; Romieu et al. 2004, 2006; Salam et al. 2007; Tung et al. 2011) and in all adults studies bar two (Alexeeff et al. 2008; Yang et al. 2005). This is followed by GSTP1 genotype which is located on chromosome 11q13, with ten studies performed in children (Islam et al. 2009; Lee et al. 2004; Li et al. 2006; Melén et al. 2008; Piacentini et al. 2010; Reddy et al. 2012; Romieu et al. 2006; Salam et al. 2007; Schroer et al. 2009; Tung et al. 2011) and nine studies in adults (Alexeeff et al. 2008; Avogbe et al. 2005; Binkova et al. 2007; Castro-Giner and et.al 2009; Chen et al. 2007b; Gilliland et al. 2004; Novotna et al. 2007; Ren et al. 2010; Topinka et al. 2007). In contrast the GSTT1 genotype on chromosome 22q11 has been examined in only 4 studies in children (Buthbumrung et al. 2008; Hong et al. 2007; Piacentini et al. 2010; Salam et al. 2007) and 8 studies in adults (Avogbe et al. 2005; Binkova et al. 2007; Castro-Giner and et.al 2009; Gilliland et al. 2004; Manini et al. 2006; Novotna et al. 2007; Ren et al. 2010; Topinka et al. 2007).

The majority of the published evidence indicates that polymorphisms in the GSTM1 genotype resulting in the homozygous deletion of GSTM1 alleles (GSTM1 null) results in a significant increased susceptibility to air pollutants in children and adults with only two studies failing to confirm this association (Li et al. 2006; Manini et al. 2006). The study by Li et al. (2006) on acute exposure of ozone demonstrated decreased risk of wheeze in subjects who were GSTM1 null, whereas in the Manini et al. (2006) study, increased risk to benzene was observed in GSTM1pos alleles particularly in smokers.

In contrast, investigation of the GSTP1 genotype has yielded mixed results. SNP Ile105Val has been of particular interest in all studies, with evidence that individuals

with homozygous Val alleles exhibit significantly lower GST enzyme activity (Watson et al. 1998). In four studies, exposure to high ozone level demonstrated substantial risk to oxidative stress biomarkers (Avogbe et al. 2005) and significant decreased lung function (Alexeeff et al. 2008; Chen et al. 2007b) and breathing difficulties (Romieu et al. 2006) for those with both allele Valine. Similarly, amongst children with high EPHX1 activity and homozygous Val/Val, living near major road had increased susceptibility to asthma (Salam et al. 2007). Conversely, a substantial protective effect against wheezing was observed amongst subjects with variant Ile/Ile (Li et al. 2006), or with either one or both allele Val (Lee et al. 2008) when exposed to low air pollution and high ozone level respectively. However, another study (Schroer et al. 2009) showed significant increased risk of wheezing for subjects with homozygous Val/Val compared to wild type with respective of pollutant exposures. In a recent study, carriers with homozygous Val/Val showed increased FEV_1 in relation to acute exposures to PM_{10} and SO_2 compared to non-carriers (Reddy et al. 2012). Conversely, possession of wild type Ile/Ile, which has been shown to be associated with increased enzyme activity, indicated enhanced susceptibility to asthma in some studies (Islam et al. 2009; Lee et al. 2004) in individuals with high exposure to air pollution. In addition, in a crossover study, enhanced nasal sensitization was observed in subjects with homozygous Ile/Ile who were exposed to DEP (Gilliland et al. 2004). However, substantial increased risk to sensitization was demonstrated for children with heterozygous Ile/Val exposed to elevated levels of traffic NOx (Melén et al. 2008). Interestingly, subjects with Val/Val demonstrated a six fold decreased risk of asthma compared to controls (Fryer et al. 2000). In these studies, it remains unclear to what extent the period of pollutant exposure contributes to the different outcome observed. In addition, how these functional polymorphisms relate to other aspects of the pulmonary antioxidant network is not well defined at this time. It is however notable that chromosome 11q13 has been shown to be hot spot for asthma related genes (Thomas et al. 1997; Ferreira et al. 2011), which has led to speculation as to whether the GSTP1 genotype may have a linkage to the atopic asthma phenotype.

Table 1.3 Studies of interactions between genetic polymorphism and air pollution in adults

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Bergamaschi et al. 2001	Italy/ White	Case control/ Blood lymphocytes/ Multiplex PCR and PCR-RFLP	Healthy subjects=24/ Males=9 Females=15/ 28.5 ± 3.4 years	Ozone= 32 and 103 ppb (median 78 ppb) (2 hour mean ozone) exposure	GSTM1- deletion NQO1-Pro187Ser	NA rs1800566	Subjects with NQO1 Pro/Pro and GSTM1 null showed significant ↓ PEF and FEV and significant↑ in serum CC16 after acute exposure to O_3. But no changes in lung function.
Corradi et al. 2002	Italy/ White	Case control/ Blood lymphocytes/ Multiplex PCR and PCR-RFLP	Healthy subjects=22/ Males-12 Females-10/ 30.1±2.6 years	Ozone exposure 0.1ppb (2hours) with moderate exercise	GSTM1- deletion NQO1-Pro187Ser	NA rs1800566	Subjects with GSTM1 null and NQO1 Pro/Pro show significant changes in EBC and ↑ of 8-OHdG when exposed to 2h O_3.
Apostoli et al. 2003	Italy/ Caucasians	Case control/ Blood lymphocytes/ PCR-RFLP	Controls=30 Exposed=171/ All males/ 18-60 years	PAH-[Pyrene] in air- Low exposure 35.5 ng/m^3 High exposure 124.0 ng/m^3 Very high exposure 316.0 ng/m^3	CYP1A1 i)CYP1A1*2A (Msp1) ii) CYP1A1*2B (Ile462Val) GSTM1-deletion	rs4646903 rs1048943 NA	No influence of the genetic polymorphism of CYP1A1 and GSTM1 on the urinary levels of 1-OH-P.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Gilliland et al. 2004	United States/ White, Hispanic, African American, Asian	Case control/ Buccal cells/ TaqMan Allelic Discrimination (AD) assay	N=19 ragweed sensitive/ Males=7 Females=12/ 20-34 years	Ragweed alone, ragweed plus 0.3mg diesel exhaust particles	GSTM1- deletion GSTT1 - deletion GSTP1-Ile105Val	NA NA rs1695	Subjects with GSTM1 null greater ↑ IgE and histamine after diesel exhaust particles plus allergen challenge. Subjects with GSTP1 Ile/Ile significant ↑ IgE and histamine after diesel exhaust particles plus allergen challenge. Similarly in subjects with both GSTM1 null and GSTP1 Ile/Ile variants had greatest ↑.
Yang et al. 2005	Germany	Case control/ Blood lymphocytes/ TaqMan Allelic Discrimination (AD) assay	N=51 undergone ozone challenge/ Males=26 Females=25/ Mean 29±6 years	Ozone: 250 ppb for 3 hours (44 subjects) 200 ppb for 4 hours (4 subjects) 400 ppb for 2 hours (3 subjects)	SOD_2 - Val16Ala GPx1- Pro197Leu	rs4880 rs1050450	No significance changes in FEV1 with SOD2 or GPx1 genotypes.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Avogbe et al. 2005	Denmark/ Danish	Case control/ Blood lymphocytes/ PCR-RFLP and TaqMan AD assay	Taxi-moto drivers: N=29 / 36±6, Roadside Residents: N=37/27±8 Suburban: N=42/ 36±13 Rural: N=27/40±12 All males	i)Ultra-fine particle (numbers/cm^3) mean±SD Rural: 6961±3223 Suburban: 19,980±10,798 City background: 41,230±5822 Intersection at Balcony: 160,194±56,908 Traffic circle: 201,691±49083 Intersection in the street: 265,145±76111 ii) Benzene City:76.0±26.8 µg/m^3 Rural: 3.4±3.0 µg/m^3	GSTM1-deletion GSTT1-deletion GPx1-Pro198Leu GSTP1 -Ile105Val NQO1-Pro187Ser	NA NA rs1050450 rs1695 rs1800566	Subjects with GSTT1 null genotype had significant ↓ urinary S-PMA excretion. Subjects with NQO1 Pro/Ser have strongest correlation between S-PMA and SB. Subjects with the GSTP1 Val/Val have strongest correlation between S-PMA and FPG sensitive sites. Stepwise gradients with respect to ambient UFP, S-PMA excretion and oxidative DNA damage with rural subjects < suburban subjects < residents living near highly trafficed roads < taxi-moto drivers.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Manini et al. 2006	Italy/ Italian	Case control/ Blood lymphocytes/ PCR-RFLP	N=37/ Males=35 Females=2/ 42.7±11.7 years	Airborne benzene=: mean±SD: 5.85±1.65µg/m^3 (24hrs personal exposure) 7.71±1.95µg/m^3 (work-shift)	GSTM1-deletion GSTT1-deletion GSTA1*A/*B i)-631(G/T) ii)-567 (T/G) iii)-69(C/T) iv)-52(G/A) NQO1-Pro187Ser EPHX1 i) Tyr113His ii) His139Arg	NA NA rs4715333 rs4715332 rs3957356 rs3957357 rs1800566 rs1051740 rs2234922	Subjects with GSTM1pos had significant ↑ excretion of S-PMA particularly in smokers group.
Binkova et al. 2007	Czech Republic	Case control/ Blood lymphocytes/ PCR-RFLP and TaqMan AD assay	Control=52 Exposed= 53/ All males/ 22–50 years	PM_{10} =32-55 µg/m_3 $PM_{2.5}$=27-38 µg/m^3 c-PAHs=18-22 ng/m^3 Personal exposure (cPAHs)= 9.7 ng/m^3 vs 5.8 ng/m^3	GSTM1-deletion GSTT1-deletion GSTP1- Ile105Val EPHX1 i) Tyr113His ii) His139Arg CYP1A1 i) Msp1 ii) Ile462Val XRCC1-Arg399Gln XPD- Arg156Arg XPD-Lys751Gln	NA NA rs1695 rs1051740 rs2234922 rs4646903 rs1048943 rs25487 rs238406 rs13181	Exposed subjects with GSTM1 null, and XPD23 (Gln/Gln) significant ↑ total DNA adduct levels. Control subjects with XPD6 (C/C) significant ↑ B[a]P-like DNA adduct levels.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Chen et al. 2007	San Fransisco, United States/ Asians, Caucasians, Hispanic	Cohort/ Blood lymphocytes/ TaqMan Allelic discrimination (AD) assay	N=210/ Males (M) =90 Females(F)= 120/ <18 year: M=46, F= 57 19 year: M=49, F= 41 >20 year: M=5, F= 2	Lifetime exposure: (Males, Females) Ozone= (3, 33)ppb PM_{10} = Prior 1987: (73, 69) µg/m^3 After 1987 (36, 29) µg/m3 NO_2 = (29,26)ppb	GSTM1- deletion GSTP1- Ile105Val NQO1-Pro187Ser	NA rs1695 rs1800566	Female subjects with GSTM1 null/ NQO1 Pro1/Pro significant ↑ risk of ozone related ↓ in mean FEF(25-75%). Male subjects with GSTP1 Val/Val significant ↑ risk of ozone related ↓ in mean FEF (25-75%).
Novotna et al. 2007	Czech Republic	Case control/ Blood lymphocytes/ PCR-RFLP	65 non-smokers/ All males/ Exposed= median 31 years Control- median 35 years	B[a]P (ng/m3): Jan- 2.65±2.86 Sept- 1.15±1.15 cPAH (ng/m3): Jan-16.6±17.9 Sept- 8.6± 7.3 $PM_{2.5}$ µg/m^3: Jan-33.2±39.7 Sept-14.5±8.6	CYP1A1 i) Msp1 ii) Ile462Val GSTM1-deletion GSTT1-deletion GSTP1 -Ile105Val EPHX1-Tyr113His XRCC1-Arg399Gln XPD- Arg156Arg XPD- Lys751Gln hOGG1-Ser326Cys	rs4646903 rs1048943 NA NA rs1695 rs1051740 rs25487 rs238406 rs13181 rs1052133	Exposed subjects with variant CYP1A1*2C (Ile/Val) and EPHX1-medium (Tyr/His) phenotype significant ↑ susceptible to the induction of oxidative DNA damage. Winter levels of oxidative DNA damage positively correlated with exposure to cPAHs, probably reflecting ↑ oxidative stress as a result of ↑ of [PM2.5].

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Topinka et al. 2007	Czech Republic	Case control/ Blood lymphocytes/ PCR-RFLP	109 non-smokers/ All males/ 35±0.9 years	B[a]P (ng/m^3): Jan-2.65 ± 2.86, Mac-0.63 ± 0.35, June-0.48 ± 0.56 Sept- 1.15 ± 1.15 cPAHs (ng/m^3): Jan- 16.6 ± 17.9 Mac- 3.7 ± 1.8 June- 3.5 ± 2.8 Sept- 8.6 ± 7.3 To cover high (winter) and low (summer) air pollution	CYP1A1 i) Msp1 ii) Ile462Val GSTM1-deletion GSTT1-deletion GSTP1 -Ile105Val XRCC1-Arg399Gln XPD- Arg156Arg XPD- Lys751Gln hOGG1-Ser326Cys	 rs4646903 rs1048943 NA NA rs1695 rs25487 rs238406 rs13181 rs1052133	Subjects with both CYP1A1 (Ile/Val) and GSTM1 null had significant ↑ total and B[a]P-like DNA adducts with ↑ exposure of cPAHs.
Alexeef et al. 2008	Boston, United States/ White & black	Cohort/ Blood lymphocytes/ Sequenom Mass Array MALDI-TOF mass spectrometer & short tandem repeat (STR) assay	N=1015/ All males/ 68.9±7.2 years	Ozone, 2-day mean, ppb 24.4 ± 11.0	GSTP1 i) Ile105Val ii) Ala114Val HMOX1 - (GT)n repeat	 rs1695 rs1138272 NA	Subjects with a long (GT)n repeat in HMOX1 or Val105 in GSTP1 had significant ↓ in FEV1 with a 15 ppb increase in O3. Subjects carrying both the GSTP1 105Val variant and the HMOX1 long (GT)n repeat had a stronger effect of O3 on FEV1.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Castro-Giner et al. 2009	Six European countries Mainly white	Cohort/ Whole blood/ PCR; Pyrosequencing; SNPlex™ platform	N=2577/ Females=52.2%/ 43.03 ± 7.3 years	NO_2	GSTM1-deletion GSTT1-deletion GSTP1-Ile105Val NQO1 i) 3′ UTR (C/T) ii) Pro187Ser iii)3′ region (C/G)	NA NA rs1695 rs10517 rs1800566 rs2917666	Subjects with NQO1 CC genotype at 3' downstream had significant association between modeled NO_2 and asthma prevalence and also with new on-set asthma.
Ren et al. 2010	Boston, United States	Cohort/ Blood lymphocytes/ Sequenom Mass Array MALDI-TOF mass spectrometer & TaqMan 5' exonuclease (ABI)	N=320/ All males/ 76.7 ± 6.1 years	Organic carbon (OC), Maximal one-hour ozone, Sulfate (SO_4^{2-})	CAT i) C/T ii) A/G iii) G/A iv) A/G GSTM1-deletion GSTT1-deletion GSTP1 i) Ile105Val ii) Ala114Val HMOX1 i) A/T ii) C/G iii) G/A iv) C/G GCLC GCLM i) A/G ii) A/G	rs480575 rs1001179 rs2284367 rs2300181 NA NA rs1695 rs1799811 rs2071746 rs2071747 rs2071749 rs5995098 rs17883901 rs2301022 rs3170633	Subjects carrying wild type GSTP1 (rs1799811) marginally significantly modified the protective effect of SO_4^{2-} on 8-OHdG (adjusted p =0.091). Subjects with the wild type of CAT (rs2284367) (adjusted p = 0.037), and the non-deletion of GSTM1 (adjusted p = 0.037) significantly modified effects of OC on 8-OHdG.

CYP1A1 = cytochrome *P*-450 enzymes; *NQO1* = nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase; glutathione *S*-transferase *(*GST*) M1, GSTP1, GSTT1, GSTA1*; *GPX1* = glutathione peroxidase; *EPHX1* = microsomal epoxide hydrolase 1; *TNF-α* = tumor necrosis factor alfa; *LTA* = lymphotoxin; *TLR4* = Toll-like receptor 4; *SOD2* = superoxide dismutase; XPD = xeroderma pigmentosum D; XRCC1 = X-ray repair cross-complementing group 1; *MS* = methionine synthase; *MTHFR* = methylenetetrahydrofolate reductase; *hOGG1* = 8-oxoguanine DNA glycosylase; Pro = proline; Ser = serine; Ile = isoleucine; Val = valine; Tyr = tyrosine; His = histidine; Arg = arginine; Ala = alanine; Gly = glycine; Gln = glutamine; Asp = aspartic acid; Leu = leucine; Cys = cysteine; A = adenine; C = cytosine; G = guanine; T = thymine; SNP = single nucleotide polymorphism; $PM_{2.5}$ = particulate matter with aerodynamic diameter less than 2.5 microns; PM_{10} = particulate matter with aerodynamic diameter less than 10 microns; NO_2 = nitrogen dioxide; ppb = parts per billion; PEF = peak expiratory flow; FEF = forced expiratory flow; FEV = ↑ = increase, high, more; ↓ = decrease, low; EBC= exhaled breath condensate; c-PAH= carcinogenic polycyclic aromatic hydrocarbons; S-PMA= S-phenylmercapturic acid ; SB= strand breaks; FPG= formamidopyrimidine DNA glycosylase; B[*a*]P = benzo[*a*]pyrene; 1-OH-P= 1-hydroxypyrene; Pos=positive; PCR-RFLP= polymerase chain reaction-restriction fragments length polymorphism.

Table 1.4 Studies of interactions between genetic polymorphism and air pollution in children related to respiratory disease

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
David et al. 2003	Mexico/ Mexican	Case triad/ Blood lymphocytes/ PCR-RFLP	N= 218/ Boys=58% Girls =42%/ 4-17 years	Lifetime exposure Ozone= mean 102±47ppb	NQO1- Pro187Ser GSTM1-deletion	rs1800566 NA	Subjects with NQO1 Pro/Ser have ↓ asthma risk only among GSTM1 null subjects in population of high ozone level exposure.
Lee et al. 2004	Taiwan/ Taiwanese	Cross-sectional/ Buccal cells/ PCR-based assay	Asthmatic=61 Boys=33 Girls=28 11.8±1.5 years Controls=91 Boys=54 Girls=41 12.1±1.8 years	Outdoor air pollution Low : Mean NOx=22.6, SO_2=3.6ppb Moderate: Mean NOx=26.2, SO_2=6.2 ppb High: Mean NOx = 41.2, SO_2=8.6ppb	GSTP1 -Ile105Val	rs1695	Subjects with Ile/Ile (in high air pollution) have significant ↑ asthma risk but subjects in low/ moderate pollution areas did not show ↑ risk.
Romieu et al. 2004	Mexico/ Mexican	Randomised controlled trial/ Blood lymphocytes/ Semi-quantitative PCR	N= 158/ Boys=102 Girls=56/ Mean 9.2 years	Ozone= mean 102±47ppb (1 h max ozone)	GSTM1-deletion	NA	Subjects receiving placebo with GSTM1 null significant ↓ FEF22-75 in 50ppb ozone. Subjects with GSTM1 null significant ↑ benefit of antioxidant supplementation.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Li et al. 2006	California, United States/ Non-Hispanic white (2,182)/ Hispanic white (989)	Cohort/ Buccal cells/ Real time PCR (allele-specific minor groove binder probes)	N= 3699/ Boys=1,721 Girls=1,978 8-18 years	Ozone Low ozone= 37.5 ppb High ozone= 57.8 ppb	TNF-α - G-308A GSTM1-deletion GSTP1- Ile105Val	rs1800629 NA rs1695	Subjects with TNF-308 GG significant ↓ asthma and lifetime wheezing risk in low ozone. Significant ↓ wheeze risk with TNF-308 GG among GSTM1 null subjects in low ozone. Significant ↓ wheeze risk with TNF-308 GG among GSTP1 Ile/Ile subjects in low ozone.
Romieu et al. 2006	Mexico/ Mexican	Randomized controlled trial/ Blood lymphocytes/ TaqMan Allelic discrimination (AD) assay	N= 151/ Boys=95 Girls=56/ Mean 9.2 years	Ozone=mean 102±47ppb (1h max ozone) PM_{10}=mean 56.68±27.36 (24h average) NO_2=mean 66±39ppb (1h max)	GSTM1-deletion GSTP1- Ile105Val	NA rs1695	Subjects with GSTM1 null or GSTP1 Val/Val significant ↑ in breathing difficulty with a 20ppb increase ozone exposure. Similarly in subjects with both GSTM1 null and GSTP1Val/Val variants.
Hong et al. 2007	Korea/ Korean	Panel/ Buccal cells/ PCR based assay	N= 43/ Boys=23 Girls=20/ Mean 9.6 ± 1.1 years	PM_{10}=35.30±23.48 $PM_{2.5}$=20.27± 8.23 Fe=0.208 ± 0.203 Mn=0.008 ± 0.005 Pb =0.051 ± 0.031 Zn =0.021 ± 0.021 Al =0.085 ± 0.100 All in $\mu g/m^3$ conc.	GSTM1-deletion GSTT1- deletion	NA NA	No significant effects of GSTM1 null and GSTT1 null on PEF.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Salam et al 2007	California, United Sates Non-Hispanic white (2,124)/ Hispanic white (940)	Cohort Buccal cells TaqMan (AD) assay	N= 3124 Boys-1,459 Girls-1665 9-15 years	Major road (<75m or ≥75m)	EPHX1- Tyr113His - His139Arg GSTP1- Ile105Val GSTM1- deletion GSTT1- deletion	rs1051740 rs2234922 rs1695 NA NA	Subjects who lived <75m with high EPHX1 activity, significant ↑ risk of lifetime asthma. Subjects who lived <75m with high EPHX1 activity and with GSTP1 Val/Val significant ↑ risk of asthma.
Buthbumrung et al. 2008	Bangkok, Thailand	Cross sectional/ Blood lymphocytes/ PCR-RFLP	Urban area=109 Rural area=62/ Only boys/ Mean 10.72±0.11 years	Benzene- roadside and inside school Urban area= mean 8.25ppb Rural area (inside school)= mean 2.71ppb	GSTM1-deletion GSTT1-deletion CYP2E1*5 NQO1- Pro187Ser GSTA1*A/*B i)-631(G/T) ii)-567 (T/G) iii)-69(C/T) iv)-52(G/A)	NA NA rs2031920 rs1800566 rs4715333 rs4715332 rs3957356 rs3957357	Subjects with GSTM1 null significant ↑ urinary muconic acid (MA) excretion of benzene effects. No significant effects in other variants with related to benzene level.
Islam et al. 2008	California, United States/ Non-Hispanic white (1,125)/ Hispanic white (576)	Cohort/ Buccal cells/ TaqMan Allelic discrimination (AD) assay	N= 1701/ Non-Hispanic white boys-519 Hispanic white boys-215/ 7-9 years 10-11 years >11 years	Ozone levels Higher ozone= mean 55.2 ppb Low ozone= mean 38.4 ppb	HMOX1 (GT)n repeat CAT i) C-262T ii) C-844T MNSOD- Val16Ala	NA rs1001179 rs769214 rs4880	Non-Hispanic white subjects with HMOX1 "s" alleles significant ↓ new on-set asthma risk especially in low ozone area. Hispanic whites subjects with variant C-262T significant ↑ risk of new on-set asthma. No significant effects between non-ozone pollutants with any other variants.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Lee et al. 2008	Taiwan/ Taiwanese	Case-control/ Oral mucosa/ PCR based assay	N=399/ Boys=206/ Mean=12 years	Ozone	GSTM1-deletion GSTP1- Ile105Val	NA rs1695	Amongst the children with GSTP1-105 Ile/Val or Val/Val and GSTM1 null, children with the G allele of FcεRIβ had protective effect of wheezing compared with the wild type genotype, in higher ozone level.
Melen et al. 2008	Sweden/ Swedish Scandinavia - 84% both parents born in Scandinavia Non-Scandinavian - 16% one or both parents born outside Scandinavia	Cohort/ Blood lymphocytes/ MALDI-TOF mass spectrometry (Sequenom)	N= 982/ Boys=522 Girls= 460/ Mean 4 years	PM_{10} =4 μg/m^3 NOx=23 μg/m^3 SO_2 =3 μg/m^3 NOx=44μg/m^3 [5th–95th percentile] difference	GSTP1 i)Intron 4 (C232+13A) ii) Exon 5- Ile105Val iii) Intron 5 (C/T) iv) Intron 5 (C/T) v) Exon 6-Ala114Val vi)Intron 6-C445-16T vii) Exon 7 – T555C TNF-α i) G-308A ii) T-1031C iii) C-857T iv) Intron 1 (C/T) v) Intron 3 (A/G) ADBR2 i) Exon 1-Glu27Gln ii) Exon 1-G252A iii) Exon 1-C523A	 rs762803 rs1695 rs749174 rs1138272 rs1871042 rs4891 rs749174 rs1800629 rs1799964 rs1799724 rs1800610 rs3093664 rs1042714 rs1042717 rs1042718	Subjects with heterozygote GSTP1 Ile105Val significant ↑ risk of sensitization when exposed to elevated level of traffic NOx. Interaction effect of GSTP1 Ile105Val with traffic NOx significant ↑ risk of sensitization among subjects TNFG-308A. No significant interactions between variants TNF and ADBR2 with respiratory symptoms, PEF or sensitization when exposed to traffic NOx.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Islam et al. 2009	California, United States/ Non-Hispanic white (1,064)/ Hispanic white (546)	Cohort/ Buccal Cells/ TaqMan Allelic discrimination (AD) assay	N= 1610/ Non-Hispanic White boys-491 Hispanic white boys-226/ 7-9 years, 10-11 years & >11 years	Ozone Higher ozone= mean 55.2 ppb Low ozone= mean 38.4 ppb	GSTP1 i)5' flanking (T/A) ii) Intron (C /G) iii) Ile105Val iv) Intron (C/T) GSTM1- deletion	rs6591255 rs4147581 rs1695 rs749174 NA	Subjects who participated ≥ 3 sports with GSTP1 105-Ile/Ile significant ↑ asthma risk in high ozone area. Subjects with GSTM1 null significant ↑ asthma risk without exposure to ozone.
Schroer et al. 2009	Cincinnati, Ohio, United States/ Caucasian (464), non-Caucasian 106 (18.6%) Of non-Caucasian 86.8% were African Americans	Cohort/ Buccal Cells/ LightCyler platform (Roche Diagnostics, GmbH, Mannheim, Germany)	N=570/ Boys=298 (52.3) Girls=272 (47.7)/ Age=12-24 months	Estimated diesel exhaust particle (DEP)= 0.34 $\mu g/m^3$ (range=0.23-0.88).	GSTP1- Ile105Val	rs1695	Subjects carrying Val/Val alleles had significant ↑ risk of wheezing in high DEP exposure.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Piacentini et al. 2010	Chieti, Pescara and Rome, Italy	Case-control/ Buccal cells/ Multiplex PCR; PCR-RFLP	N=131/ Male:Female ratio- Asthmatics: 0.42 Controls: 0.40 Age asthmatics: 47.53 ± 2.05/ Age controls: 50.70 ± 2.01	PM_{10} O_3 NO_2 SO_2	GSTM1-deletion GSTT1- deletion GSTA1 (–69C/T) GSTO1- Ala140Asp GSTO1-Glu155 deletion GSTO2- Asn142Asp GSTP1- Ile105Val	NA NA rs3957356 rs4925 rs11509437 rs156697 rs1695	Subjects with GSTA1 –69T/–69T and heterozygous genotype had ↑ risk of asthma development, which was 7.15-fold higher than the risk in those with the homozygote wild type genotype. Subjects with GSTM1 null had ↑ risk of asthma development, which was 2.34-fold higher than the risk in those with the GSTM1 positive genotype. Subjects with GSTO2 homozygous Asp/Asp genotype had ↑ risk of asthma development, which was 9.10-fold higher than the risk in those with the GSTO2 positive phenotype. Subjects with GSTT1 null genotype had a 3.61-fold ↑ risk of asthma than those with the GSTT1 positive phenotype.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Breton et al. 2011	California United States/ Non-Hispanic white (1397)/ Hispanic white (709)	Cohort/ Buccal Cells/ GoldenGate platform (Illumina Inc, San Diego, CA)	N= 2,106/ Boys=1022 (48.5)/ Age= 10.0 ± 0.6	NO_2 $PM_{2.5}$ PM_{10} O_3 OC Elemental carbon	GSS	rs6087649 rs1801310 rs2273684 rs6060124 rs6060127 rs3761144 rs3761143	Subjects carrying ''0100000'' haplotype in GSS, had negative effects of all the pollutants except ozone with significant ↓ in average growth of FEV1, FVC, and MMEF over an 8-year period.
					GSR	rs12707730 rs3594 rs1465477 rs2551715 rs8190996 rs4628224 rs3779647 rs2978665 rs2978663 rs1125853 rs3779648 rs2161850 rs8190896 rs2251780 rs1002149 rs12543000 rs11986256 rs12543324	The magnitude of effects within this haplotype ranged from 2124.2 to –149.1 for FEV1, from 292.9 to 2126.7 for FVC, and from 2193.9 to –277.9 for MMEF for all pollutants except ozone. *SNP haplotypes (0100000) were coded as "1" for variant and "0" for common variant allele. They were ordered as follows: rs6087649, rs1801310, rs2273684, rs6060124, rs6060127, rs3761144, and rs3761143.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Breton et al. 2011 (continue)					GCLM	rs10874809	
						rs2064764	
						rs769211	
						rs7515191	
						rs3789453	
						rs1473711	
						rs743111	
						rs6699912	
						rs990576	
						rs6696758	
					GCLC	rs1901773	
						rs2066511	
						rs2277108	
						rs1014852	
						rs642429	
						rs3736729	
						rs84933	
						rs572494	
						rs2066508	
						rs1002269	
						rs3799695	
						rs6458941	
						rs510088	
						rs4712036	

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Tung et al. 2011	Taiwan/ Taiwanese	Cohort/ Buccal cell/ TaqMan Allelic discrimination (AD) assay	N=3,741/ Boys=1,840 (49.2) Girls= 1,901 (50.8)/ Age= 12.8 ± 0.4	NO_2	EPHX1 i)Exon 3 Tyr113His ii)Exon 4 His139Arg GSTP1 - Ile105Val GSTM1 - deletion	 rs1051740 rs2234922 rs1695 NA	Children carrying Arg/His or Arg/Arg genotypes at EPHX1 His139Arg locus, were significantly associated with significant ↑ risk of lifetime asthma and early-onset asthma (adjusted OR [aOR] = 1.3; 95% CI, 1.1-1.7; and aOR = 1.5; 95% CI, 1.1-2.1, respectively). Children live in higher NO_2 communities with EPHX1 139Arg and 113Tyr-139Arg alleles have greater risk compared those live in lower NO_2 communities. This observation was most marked within children with GSTP1 Val allele and GSTM1 positive genotype.
Reddy et.al 2012	Durban, South Africa	Cohort/ Whole Blood/ Multiplex PCR; Taqman1 SNP Genotyping Assays	N=129/ Boys=45 (50.8) Girls= 84 (65.1)/ Age= 10.6 ± 1.1	SO_2 NO_2 NO_2 PM_{10}	GSTM1-deletion GSTP1 - Ile105Val	NA rs1695	Both GSTM1 and GSTP1 genotypes did not show any significant association with FEV1 intraday variability in crude analysis. Children carrying allele Val of GSTP1 show significant ↑ FEV1 variability of 3 days prior 24-hr average PM_{10} and SO_2 exposure.

Author	Place of study/ Ethnicity	Study type/ Sample type/ Method of genotyping	Subjects (N)/ Gender/ Age	Exposure metrics	Gene/Variants	SNP ID	Findings
Ungvari et al. 2012	Hungary/ Caucasian	Case–control/ Whole blood/ Sequenom iPLEX Gold MassARRAY technology	Asthma=307 Boys=200 Girls=107/ Age=10.55±4.74/ Controls=344 Boys=173 Girls=171 Age=21.77±13.88	NO_2	Nrf2 T/G 3′ UTR C/T 3′ UTR C/T Intron C/T Intron G/T Promoter C/T Promoter T/C Promoter T/G Promoter	 rs2588882 rs2706110 rs10183914 rs1806649 rs6721961 rs6706649 rs35652124 rs2364725	In case-only analysis revealed significant differences for combined genotype of SNPs rs2588882 and rs6721961 between asthma patients that lived in modestly or highly polluted environment (OR=0.43 (0.23–0.82), p=0.01, and OR=0.51, p=0.02, respectively).

NQO1 = nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase 1; glutathione *S*-transferase *(*GST*) M1*, *GSTP1*, *GSTT1*, *GSTA1*; *TNF-α* = tumor necrosis factor alfa; *GCLC* = glutamate cysteine ligase catalytic subunit*; GCLM* = glutamate cysteine ligase modifier subunit*; GSR* = glutathione reductase*; GSS* = GSH synthetase*; EPHX1* = microsomal epoxide hydrolase 1; *TGF-ß1* = transforming growth factor beta 1; *ADBR2* = beta 2-adrenergic receptor; *CYP2E1* = cytochrome *P*-450 enzymes 1; CAT = catalase; *HMOX1* = heme oxygenase 1; *MnSOD* = manganese superoxide dismutase; Pro = proline; Ser = serine; Ile = isoleucine; Val = valine; Tyr = tyrosine; His = histidine; Arg = arginine; Ala = alanine; A = adenine; C = cytosine; G = guanine; T = thymine; SNP = single nucleotide polymorphism; $PM_{2.5}$ = particulate matter with aerodynamic diameter less than 2.5 microns; PM_{10} = particulate matter with aerodynamic diameter less than 10 microns; NO_x = total NO + NO_2; NO = nitrogen oxide; NO_2 = nitrogen dioxide; ppb = parts per billion; Fe = ferum; Mn = manganese; Pb = plumbum; Zn = zinc; Al = aluminium; PEF = peak expiratory flow; ↑ = increase; ↓ = decrease. Phase I enzyme: *CYP2E1.* Phase II enzymes: *EPHX1*; *GSTM1*; *GSTP1*; *GSTT1*; *GSTA1*; *GSR; GSS, GCLC, GCLM; MNSOD; Nrf2; NQO1*. Inflammatory mediators: *TNF-α*; *ADBR2; TGF-ß1.*

Whilst the GSTT1 genotype has been widely studied, deletion of this gene has shown no relation to air-pollutant related respiratory symptoms in all but two studies (Avogbe et al. 2005; Piacentini et al. 2010). Notably, different GSTs metabolize distinct substrates with some overlap. GSTM1 is more sensitive to the quinone metabolites of catecholamines (dopachrome) while GSTP1 detoxifies lipid and DNA oxidation products (Fryer et al. 2000). The important substrates for GSTT1, for an example, ethylene oxide, does not appear to be present in ambient PM. Additionally, GSTP1 enzyme accounts for more than 90% of GST activity in the lung, while GSTM1 has a much more restricted expression profile. GSTT1 is least expressed enzyme in the lung instead (Hayes and Strange 2000). Consequently, the GSTT1 genotype may not be a good candidate gene for genetic associations study related to ambient PM induced respiratory health effects.

The number of publications investigating the NQO1 genotype is still limited. Only 2 studies have examined this genotype in children, whereas 7 studies have been conducted in adults. One functional polymorphism in NQO1 results in a substitution at amino acid 187, from proline to serine, resulting in decreased enzyme activity. The enzyme activity of NQO1 in subjects decreased as follows: homozygous Ser/Ser < heterozygous Pro/Ser < wild type Pro/Pro (Nebert et al. 2002). The NQO1 enzyme catalyzes reduction of quinones directly to hydroquinones, avoiding the generation of intermediate semiquinone radical. Subjects with NQO1 Pro/Pro and GSTM1 deletion have been shown to have an increased sensitivity to oxidative stress (Bergamaschi et al. 2001; Corradi et al. 2002) and reduced lung function (Chen et al. 2007) with acute and chronic exposure to ozone respectively. In contrast, the NQO1 Pro/Ser genotype has been shown to be protective against asthma risk in children exhibiting the GSTM1 null genotype in population of high ozone level exposure (David et al. 2003), whilst another study showed significant correlation with oxidative damage biomarker (Avogbe et al. 2005) for subjects living in high level of benzene and ultrafine particles.

Microsomal epoxide hydrolase (EPHX1) is a phase II enzyme, which detoxifies reactive epoxides from activated PAHs to water soluble trans-dihydrodiols. There is considerable evidence that polymorphisms in EPHX1 are associated with COPD (Hu et al. 2008). However, there is limited research related to air pollution, with only 2 studies in children (Salam et al. 2007; Tung et al. 2011) and 3 studies in adults (Manini et al. 2006; Binkova et al. 2007; Novotna et al. 2007). Mainly, 2 SNPs have been examined in EPHX1, the functional polymorphisms Tyr113His at exon 3 and His139Arg at exon 4. An *in vitro* study has reported that subjects with His113 have 50% reduced enzymatic activity whereas subjects with Arg139 have an increased (25%) enzyme activity. Phenotypic classification of EPHX1 is based on polymorphisms in both exon 3 and exon 4 that exhibit fast (no mutation in exon 3 and at least one mutation in exon 4), medium (heterozygous in exon 3 and homozygous at exon 4) and low (homozygous mutation in exon 3 and no mutation in exon 4) enzyme activity (Hassett et al. 1994). One study performed on the CHS (Children's Health Study) cohort provided a clear picture of the association between EPHX1 with traffic emissions. In this study, children with high EPHX1 activity living in proximity to major roads (<75m) had a significantly increased risk of lifetime asthma. In addition, in combination with GSTP1 105Val, which has a 7-fold higher efficiency for diol epoxides than 105Ile (Hayes and Strange 2000), the increased risk of asthma amongst the children was further enhanced. Metabolism of PAHs to the metabolites producing harmless conjugates depends mainly on the activity of GST and EPHX1 enzymes. Here the biological explanation could be that GSTP1 105Val which has reduced enzymatic catalytic activity leads to accumulation of PAH epoxides. Subsequently, the PAH epoxides are metabolized by EPHX1 with increased conversion of the PAH epoxide to trans-dihydrodiol which undergoes redox reaction with the production of ROS.

Four studies in adults have investigated the modulating effects of CYP1A1 polymorphisms on population responses to air pollutants however; none of the studies was investigated in relation to respiratory disease or symptoms. In addition, only 2 studies (Novotna et al. 2007; Topinka et al. 2007) have shown positive associations, with enhanced susceptibility only apparent under high PAH exposures. The two most

common SNPs investigated in this gene are in the restriction site (T/C) at 3' flanking region and the Ile462Val in exon 7. The polymorphism Ile to Val enhances the microsomal catalytic activity of CYP1A1.

Polymorphisms in TNF-α, have provided inconsistent results for its influence on air pollution-induced health responses. The common SNP at the 5' untranslated region of the gene, G-308A has been studied in greatest detail because of its increased secretion and promoter activity (Wilson et al. 1997). In the study of Melen et al. (2008), a significant increased risk of allergic sensitization was observed in subjects with TNFG-308A in combination with the GSTP1 Ile105Val genotype. Conversely, TNF-308G/G showed a protective effect on asthma prevalence and lifetime wheeze in children with GSTP1 Ile/Ile who live in areas of low ozone. Previously, genetic association studies had shown significant correlation of TNFG-308A with association between NO_2 and asthma prevalence and new onset asthma (Castro-Giner et al. 2008).

1.4 Project Aims

This thesis includes a series of studies examining the responses of children's lungs to traffic pollutants and explores the basis for differences in individual susceptibility to these gaseous and particulate agents. ***The overarching hypothesis of the study was that the introduction of London's Low Emission Zone would be associated with year-on-year improvements in the respiratory health of cross sectional panels of 8-9 year old children living within inner city London boroughs.*** Following the delay in the implementation of the third phase of the scheme from October 2010 to January 2012, my aim was somewhat modified to focus solely on the first three years of the scheme as a baseline case for the later intervention in 2010. Nevertheless, one might still expect to see year-on-year improvements in air quality as a function of the modernization of the fleet, partly driven by the existence of the scheme. The study therefore aimed to address the relationship between traffic pollutant exposures with respiratory symptoms (chapter 3) and measures of lung function (chapter 4) over the first three years of the schemes operation.

As a secondary aim I investigated how polymorphisms in selected genes (reflecting toxicologically relevant pathways) impacted on the observed associations between children's lung function and measures of oxidative stress with their estimated chronic, sub-chronic and acute exposures to NOx, NO_2, PM_{10} and $PM_{2.5}$. This work described in chapters 4 (lung function) and 5 (urinary measures of oxidative stress and metal exposure biomarkers) is based on a separate cross sectional panels of children, (school year 4, UK) sampled during the winter periods (Nov-Mar) of 2008-2009, 2009-2010 and 2010-2011. This adjunct work focused on SNPs within key antioxidant and xenobiotic candidate genes. To examine the possible role of diesel derived polyaromatic hydrocarbons exposures on long term lung growth I examined SNPs in key phase I and II xenobiotic metabolism genes: glutathione S transferases (GST), NAD(P)H quinone oxidoreductase (NQO1), cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1), the aryl hydrocarbon receptor (AhR). For potential effect modification of the association between systemic (urinary) markers of oxidative stress and air pollutant exposures I focused on SNPs within glutamate-cysteine ligase (GCLM (modifier subunit) and GCLC (catalytic subunit), the nuclear factor (erythroid-derived 2)-like 2 transcription factor (Nrf2) and extracellular superoxide dismutase (EC-SOD/SOD3), as each of these genes has been shown to play a significant role in protecting the lung from oxidative stress.

In the final experimental chapter I have also presented preliminary work examining the potential utility of using urinary metals as biomarkers of personal exposure to traffic pollution. Here I have hypothesised that increased exposure to traffic pollution would be associated with increased urinary metals reflecting different vehicular sources: tailpipe emissions (Ni, Cr and V) and tire / brake wear (Ba, Cu and Sb).

Chapter 2

Materials and Methods

2.1 Introduction

This methods section will focus on the saliva sampling performed on the children attending schools in East London and the subsequent genotyping of susceptible genes to air pollutants. These methods are therefore relevant to the data presented in chapters 3, 4 and 5. The experimental methods employed to optimize DNA extraction and quality control will be covered in detail here. For context however, a brief description will also be provided of the additional measurements being performed on the participating children. Detailed information on lung function testing is discussed in chapter 3 and urinary analysis for markers of oxidative stress and metal exposure biomarkers will be discussed in the chapter 5. As the work in this thesis represents part of a large epidemiological study, it was dependent on the teamwork and collaboration between number individuals across various institutions. Lung function measurements were performed by trained staff from Queen's Mary hospital under the supervision of Dr Isobel Dundas, who was also responsible for quality assurance and quality control of the data. Urinary measurements of cotinine, creatinine and 8-isoprostane were performed by Dr Helen Wood of the MRC-HPA Centre for Environment and Health (chapter 5), who was also responsible for the maintenance of the master data base. Urinary 8-oxodG and metals determinations (chapter 5) were performed by Ellie Smith, an MSc student under my supervision, with Andy Cakebread of the Mass Spectroscopy unit at King's College London providing expert guidance in the use of inductively coupled plasma mass spectroscopy (ICP-MS). Dr Nadine Marlin of the Pragmatic Clinical Trials Unit, Queen's Mary University, performed statistical analysis of the data.

2.2 London Low Emission Zone Schools Study Design

We carried out yearly cross sectional surveys of 8-9 year old children's respiratory health from 2008 onwards. We invited primary schools in the east London boroughs of Hackney and Tower Hamlets to take part in the study based on their proximity to major roads and LEZ indicator air pollutant monitoring sites. We visited schools to carry out assessments during November to March each year from 2008 onwards. In this thesis, I will be concentrating on analysis performed on the first three years of the study, covering the winter periods 2008-2009, 2009-2010 and 2010-2011. Children's respiratory health was assessed by spirometry (Microlab, Micromedical, Cardinal Health). Spirometry was performed according to American Thoracic Society guidelines (Pellegrino et al. 2005) with pre- and post-bronchodilator measurements, following salbutamol 400 µg administration by large volume spacer. Other assessments included: fraction of exhaled nitric oxide, FeNO (Niox Mino, Aerocrine), urinary cotinine, heavy metals, 8-isoprostane and 8-Oxo-2'-deoxyguanosine, as well as completion of an International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire and the collection of saliva to gather DNA for genotyping. Health assessment data was linked to modeled air quality data supplied by the Environmental Research Group, Modeling team at King's College London to give individual pollutant exposures for each child's school and residential address.

2.3 Consent and Ethics Approval

Approval of the study was obtained by regional ethics committees (East London & The City HA Local Research Ethics Committee 2, REC Ref Number 08-H0704-139) and conducted according to principles of the Declaration of Helsinki. Written informed consent by parent/guardian and verbal consent from children were obtained prior to any investigation.

2.4 Subject Recruitment

A different cross-sectional sample of children was examined each winter over a 3-year period (2008-2012) in the London boroughs of Tower Hamlets and Hackney as the modeled impacts of the London LEZ had been predicted to be greatest in these areas, and detailed air pollution and traffic monitoring has been put in place to assess improvements in air quality following the introduction of this traffic amendment scheme (Kelly et al. 2011). Primary schools were selected that were either in close proximity to, or some distance from, a major roads (see **Figure A1** in **Appendix A**). The study involved 8-9 years old, Year 4 children. The parents of the children from the participating schools were provided with study packs that contained detailed information about the study and the investigations to be carried out on the pupils and the parental and children's consents were obtained prior to the investigation.

In total, 199 children from 10 schools (Nov 2008 – Feb 2009), 416 children from 19 schools (Nov 2009 – Feb 2010) and 404 children from 23 schools (Nov 2010 – Mac 2011) were sampled during the school visits for Year 1-3 which equals to 1019 subjects. The distribution of the children participation by schools is presented in the **Table A1** in **Appendix A**, whereas **Table A2** in **Appendix A** summarized details of the sampling data obtained during each visit.

2.5 Calculation of Body Mass Index

The weight and height of the child without shoes was measured in kilograms (kg), and centimeters (cm) respectively using a Stadiometer (Seca Ltd, Birmingham, UK). The height and weight were measured to determine body mass index (BMI) which was calculated using following formula: BMI = Weight (kilograms)/Height (meter)^2.

2.6 Measurement of Lung Function

Forced vital capacity (FVC) and forced expiratory volume in 1 second (FEV_1) was determined for each subject using a Spiroflow Model 132 spirometer (P K Morgan; Gillingham, UK). A minimum of three procedures were performed on each subject in accordance with the American Thoracic Society guidelines (Pellegrino et al. 2005). The measurements were obtained pre and post inhalation of four puffs of Salbutamol, to examine airway reversibility using a large volume volumetric spacer. Details on lung function QA/QC are presented in chapter 4.

2.7 Urinary Analysis

Urine samples were obtained from each of the participants on the morning of the school visit. It was not feasible to collect first morning void samples due to compliance issues. Once collected, urine samples were kept on ice until transported to King's College London where they were split into 5 – 7 of 1 ml aliquots on the same day prior to longer term storage at -80^{o}C. Samples were aliquoted for the following assays: quantification of urinary 8-isoprostane and 8-hydroxy-2'deoxyguanisine as markers of systemic oxidative stress and metal analysis as biomarker of exposure, which are covered in detail in chapter 5. In addition, urinary cotinine were to provide information on the child's exposure to environmental cigarette smoke and creatinine, as a denominator of urine concentration both of which are reported in this chapter. However, the data of these two markers was not reported, instead used for the normalization of 8-isoprostane and 8-hydroxy-2'deoxyguanisine concentrations in all statistical models.

2.7.1 Analysis of Cotinine

Cotinine is a major nicotine metabolite, with a considerably longer half-life than nicotine (about 20 h versus 2 h). It can be detected in blood, saliva and urine for up to around 72 h after exposure to nicotine. It is the biomarker of choice for quantifying

exposure to environmental tobacco smoke (ETS, i.e. passive smoking), since it is more specific for tobacco smoke than other biomarkers, such as carbon monoxide and carboxyhaemoglobin. Cotinine concentrations in urine are on average 10 times those of serum; hence urine provides a better indicator of ETS exposure. Environmental tobacco smoke (ETS) refers to a mixture of side-stream smoke and exhaled mainstream smoke that pollutes the air in locations where tobacco is being smoked. Urinary cotinine was analysed by enzyme immunoassay (EIA). A micro plate EIA kit for serum (product no. M155B1, Concateno, Abingdon, UK) with urine standards were used since the urine-specific kit is designed to measure direct tobacco exposure, i.e. from smoking, so the range of detection (50-5000ng/ml) is much higher than the levels expected in our samples (0-100 ng/ml). Farber et al. (2008) described using a cotinine saliva EIA kit to determine ETS exposure from urine samples (Chest 2008; 133:1367-1374). Cotinine levels are typically about six times greater in urine than in serum or saliva (Benowitz 1999), hence we based our method on that used by Farber et al. (2008). Undiluted urine samples were analysed in duplicate, as per the manufacturer's instructions. The kit is semi-quantitative, indicating a positive or negative result according to a cut-off point of 25 ng/ml (>25 ng/ml is positive). The measurement was repeated if the duplicate readings fell on either side of this cut-off point, or if the readings fell on the same side of the cut-off point, but with a large SD. Individual variation in urine concentration were corrected by using the cotinine-to-creatinine ratio (CCR) in ng/mg. A positive sample was defined as having a CCR > 30 ng/mg (Henderson et al. 1989). A typical standard curve for cotinine assay is illustrated in **Figure 2.1**.

2.7.2 Analysis of Creatinine

Creatinine is formed *in vivo* at a fairly constant rate, by the hydrolysis of creatine and phosphocreatine, and is cleared from the body through the kidneys, mainly by glomerular filtration. In the absence of renal disease, the excretion rate of creatinine into the urine is relatively constant (which makes urinary creatinine concentration inversely proportional to urine flow rate). Thus, urinary creatinine levels may be used as

an index of standardization for other tests. Creatinine normalization compensates for differences in urine concentration (i.e. diuresis) and hence, in biological monitoring, many compounds are expressed relative to the level of creatinine, permitting comparison between subjects and within subjects over time. Creatinine adjustment involves dividing the analyte concentration by the creatinine concentration. Urinary creatinine was analysed using a commercial kit based on the Jaffe reaction (product no. 500701, Cayman Chemical, Ann Arbor, MI, USA), wherein a yellow/orange colour forms when the metabolite is treated with alkaline picrate. The colour derived from creatinine is then destroyed at acidic pH. The difference in colour intensity measured at 500 nm before and after acidification is proportional to the creatinine concentration. Urine samples, diluted 20x were analysed in duplicate. The dynamic range of the kit is 0-15mg/dL, with an intra-assay CV of 2.7% and an inter-assay CV of 3%. The measurement was repeated if the CV for duplicates was >10%. A typical standard curve for creatinine assay is illustrated in **Figure 2.2**.

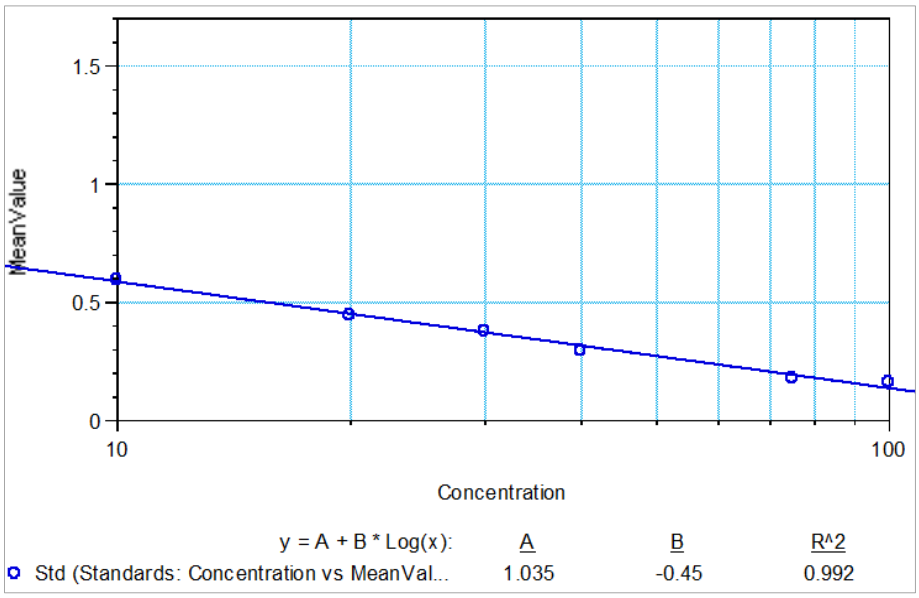


Figure 2.1 Representative of cotinine assay standard curve. Absorbance was plotted standard urine concentration.

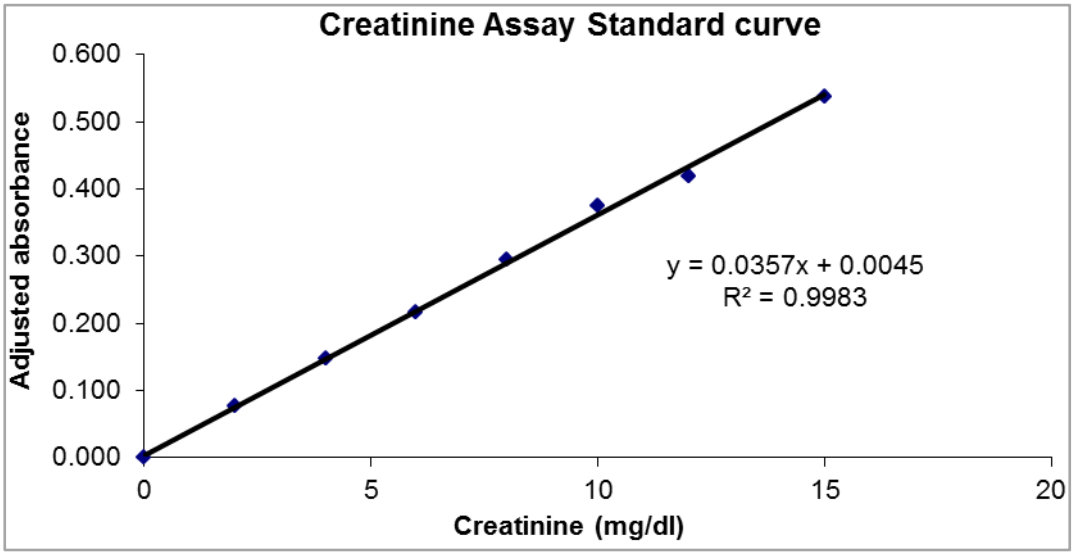


Figure 2.2 Representative of typical creatinine assay standard curve. Absorbance was plotted against standard creatinine concentration.

2.8 Saliva Collection for SNP Analysis

Saliva (a mixed saliva and buccal cell sample) was collected from children at the participating schools using the Oragene DNA kit OG-250 (DNA Genotek Inc, Canada). The collection was carried out according to the manufacturer's instruction. Briefly, the Oragene sampling sponge was placed into the childs' mouth and moved around the cheek pouches. Five sponges were performed per individual, with the saliva laden tips cut from their handle and placed in the supplied Oragene - DNA kit container. The containers were then capped with the release of a proprietary DNA stabilization solution (designed to preserve DNA in saliva samples at room temperature) and gently mixed by inverting five times. Once collected, the collection kit containing saliva samples were kept on ice until transported to King's College London. Subsequently, all samples were stored at -80^{o}C until required for processing. In addition, a preliminary study was also performed to optimize DNA recovery from two different types of saliva collection: saliva spit and saliva swabs. The preliminary study consisted of 10 volunteers, each

providing a sample of each type of collection. Details of preliminary work on DNA extraction and quantification optimization are summarized in **Appendix B**.

2.8.1 DNA Recovery from Saliva Samples

Genomic DNA was recovered from the saliva swabs according to the following protocol. After removal from storage at -80^{o}C the sample container was placed in an air incubator overnight at 50^{o}C. This incubation step was to permit cell lysis and to ensure that the DNA was released from the matrix and that nucleases are permanently inactivated. After incubation, the maximum volume of free liquid was removed and the sample transferred to a 15 ml conical centrifuge tube (Greiner, UK). The barrel of a 5 ml disposable plastic syringe was then placed into conical tube (without the plunger) into which the sponges were then transferred using fine sterile forceps. The conical tube containing sponges in the barrel was then centrifuged at 200 x g for 10 minutes at 20^{o}C. After centrifugation, the syringe barrel containing dry sponges was removed and discarded according to Human Tissue Authority (HTA) regulations.

2.8.2 DNA Extraction

Following incubation and extraction, samples were transferred to several 2.0 ml micro-centrifuge tubes with the total volume recorded. For RNase digestion, each sample was then treated with Ribonuclease A (final concentration 10 μg/ml) and Ribonuclease T1 (final concentration 25 Units/ml) both from Sigma Chemical Company Ltd (Poole, UK) and incubated at 37^{o}C for 30 minutes. Subsequently, the sample was treated with a 1/25th volume of proprietary DNA purifier (supplied as part of the Oragene DNA kit), vortexed briefly prior to 10 minute incubation on ice. This protein precipitation step resulted in a cloudy solution which was centrifuged at 13,000 rpm for 15 minutes with the resultant clear supernatant transferred to a fresh 1.5 ml centrifuge tube. An equal volume of 100% ethanol was then added and the sample gently mixed by inverting 10 times to obtain a homogenous solution. The sample was

then allowed to stand at room temperature for 10 minutes to permit complete precipitation of DNA prior to centrifugation at room temperature at 13,000 rpm for 10 minutes. The supernatant was discarded prior to the addition of 1 ml 70% ethanol to the DNA pellet and allowed to stand for 1 min at room temperature. The dehydrated DNA was then rehydrated by the addition of 100 µl of reduced TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) from Invitrogen Technologies (Paisley, UK) followed by a 30 second vortex prior to incubating for 1 hour at 50^{o}C with occasional mixing or an overnight incubation at room temperature.

2.8.3 DNA Quantification

DNA can be quantified either by UV absorbance or fluorescence-based techniques. However, the presence of impurities in saliva may cause interference in the UV spectra leading to increased absorbances at 280 nm, thereby lowering the A_{260}/A_{280} ratio. Fluorescence based methods are less subject to such variations due to the specific binding of selected fluorophores to double stranded DNA. In this study, DNA yields were determined using both UV absorbance and also fluorescence for comparison of accuracy and purity of DNA extracted from saliva.

2.8.3.1 Determination of DNA by UV Absorbance

DNA quantification by absorbance was performed using a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies, Wilmington, DE) in 1.5 µl samples. The samples were measured at 260nm to obtain an estimate of the concentration of nucleic acid and 280nm for the protein content. Generally an A_{260}/A_{280} ratio >1.8 is accepted as indicator of limited protein and organic contamination. Similarly, a ratio of A_{260}/A_{230} in the range of 1.8-2.0 illustrates limited salt, carbohydrates, ethanol and protein contamination in the sample. Samples were quantified before and after the dilution.

2.8.3.2 Determination of DNA by Fluorescence

Quantification of genomic DNA by using fluorescence was carried out by using Quant-iT™ PicoGreen® assay (Invitrogen). An aqueous working solution of Quant-iT™ PicoGreen® reagent (200 fold dilutions) was prepared in TE buffer in a 50 ml conical tube, covered with foil just few hours prior to sample analysis, as the solution is prone to photodegradation. For the DNA standard curve, the lambda DNA standard provided at 100 µg/ml in the Quant-iT™ PicoGreen® kits was diluted by serial dilution to achieve a concentration range between 0.03 and 1 ng/µl (**Table 2.1**). For analysis, each sample was diluted 1:200 with TE buffer. Subsequently, 100 µl of each of the lambda DNA standards (1-6) and samples were added in triplicate to the wells of a black 96-well plate, to which 100 µl of the Picogreen working solution of reagent was added and incubated for 2 to 5 minutes at room temperature in the dark. Following the incubation, sample fluorescence was measured on a fluorescence microplate reader (Cary Eclipse) at an excitation of 480 nm and emission 520 nm. DNA concentrations were calculated with reference to the standard curve (**Figure 2.3**) having first subtracted the background fluorescence of the blank. Details of the DNA yield across the samples can be viewed in **Table C1, Appendix C**.

Table 2.1 Protocol for preparing a standard curve for Picogreen Assay

Description	Volume of TE (µl)	Volume of lambda (λ) DNA Stock/Diluted (µl)	Final [DNA] in Quant-iT™ PicoGreen® assay (ng/µl)
Blank	100	0	0
Standard 1	1960	40 µl of 100ng/µl of DNA Stock	1
Standard 2	1000	1000 µl of Standard 1	0.5
Standard 3	1000	1000 µl of Standard 2	0.25
Standard 4	1000	1000 µl of Standard 3	0.125
Standard 5	1000	1000 µl of Standard 4	0.0625
Standard 6	1000	1000 µl of Standard 5	0.03125

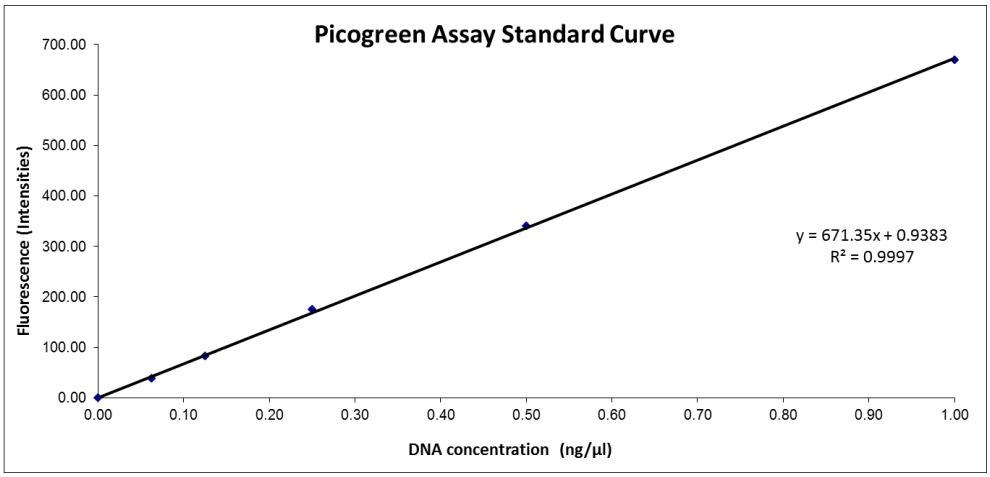


Figure 2.3 Representative picogreen assay standard curve. Fluorescence emission intensity was plotted against standard DNA concentration.

2.8.4 Genomic DNA Analysis

DNA integrity was analyzed by agarose gel electrophoresis. Briefly, 100 ng of each sample was loaded and separated on a 0.75% gel containing ethidium bromide (0.5 µg/ml) for 60 minutes at 90 volts. An instrument Syngene G:BOX was used to photograph the gel. Lambda DNA digested with HindIII (Sigma Aldrich) was used as a size reference (see **Figure C1** in **Appendix C**).

2.8.5 Whole Genome Amplification

High-throughput genotyping demands considerable quantities of genomic DNA. Therefore, due to the limited availability of the genomic DNA sources for the genotyping in certain samples, the genomic DNA present in the samples was amplified by using Whole genome amplification (WGA) technique, to permit genotyping on the Illumina platform. Although, WGA is a good technique to enhance DNA concentrations, it may also decrease the hybridization capacity of the amplified DNA. Considering that most genome-wide association studies generally involve concurrent genotyping of a huge amount of samples, this may lead to a modification in the positions of a fraction of SNPs in the genotype clusters, producing unclear clusters and

enhancing ambiguity of a call (Teo et al. 2009). Therefore, for the consistency of the study design, WGA by using the GenomePlex Complete Whole Genome Amplification Kit (WGA2) (Sigma Aldrich) was carried out as previously described (Arneson et al. 2008) on all collected samples. This amplification kit was employed as it has been tested in-house by Illumina to work in conjunction with the GoldenGate genotyping technology (Barker et al. 2004).

To begin with, the genomic DNA was normalized to 1 ng/µl. Subsequently, 1 µl of 10x DNA fragmentation buffer was added to 10 µl gDNA (1 ng/µl) in a 0.2 ml multiwell strip. The mixture was then heat-denatured at 95°C for exactly 4 min for random fragmentation and immediately cooled on ice. After a brief vortex and centrifugation to consolidate the products, 2 µl of 1 x library preparation buffer that contains degenerated adapters**,** and 1µl of library stabilization solution were added to the samples. Following that, the mixture was incubated at 95°C for 2 min and then immediately cooled on ice.

For Omniplex library construction (**Figure 2.4**), a standard PCR protocol was employed. Upon addition of 1 µl of Library Preparation Enzyme, the mixture was vortexed and centrifuged briefly. Then PCR reactions were performed in a thermal cycler as follows: 16°C for 20 min, 24°C for 20 min and finally at 37°C for 20 min. Then**,** the final extension step was performed at 75°C for 5 min. Each batch of amplification included a negative (no template DNA) control to identify non-specific amplification in the reaction and positive control (human gDNA) for the confirmation of negative amplification. The final step in the WGA was library amplification. This was accomplished via addition of a 60 µl of master mix containing 7.5 µl of 10X Amplification Master Mix, 47.5 µl of nuclease free dH_2O and 5 µl of WGA DNA Polymerase to the sample. The mixture was then vortexed and centrifuged briefly. After an initial denaturation at 95°C for 3 min, a standard PCR reaction was followed: 14 cycles of 95°C for 15s and annealing/extension at 65°C for 5 min. The PCR product was stored at -20°C. Whole genome amplification (WGA) was successfully carried out on all the DNA samples to achieve at least 50ng/µl, which is the minimum requirement for genotyping on the Illumina platform.

2.8.6 PCR Product Clean-up

The amplified genomic DNA was purified by using a PCR clean-up kit (GenElute PCR Clean-Up Kit (Sigma Aldrich)) to eliminate unincorporated primers, excess nucleotides, salts and enzymes**,** which is crucial for the success of the downstream applications. This kit employed positive selection purification and all steps were performed at room temperature at 16,000 x g unless specified. Upon inserting the GenElute Miniprep Binding Column into collection tube, 0.5 ml Column Preparation Solution was added into each column and centrifuged at 12,000 x g for 30 seconds. The eluate was discarded. Subsequently, 5 volumes of binding solution was added to 1 volume of PCR product and mixed. The solution was then transferred into binding column and spun for 1 min. Subsequently, the eluent was discarded and the bound DNA was washed with 0.5 ml Wash Solution (70% ethanol) and centrifuged for 1 min. The elutant was again discarded and bound DNA centrifuged again for 2 min without additional wash solution for complete removal of excess ethanol. The collection tube was then discarded together with the elutant and a clean 2.0 ml collection tube was replaced. Finally, 30-50 µl TE buffer (Sigma Aldrich) was added to the column and the bound DNA was eluted by centrifuging for 1 min. The purified genomic DNA was stored at -20°C until required for further analysis.

2.8.7 SNP Genotyping

It is not feasible or cost effectible to investigate large number of genes, or genetic variants using the conventional methods such as PCR. The bead-based multiplexed SNP assay has made it possible to resolve millions of SNPs in a single experiment. The Illumina GoldenGate Genotyping Assay, which uses single-base extension to call the SNP, was employed for SNPs genotyping in this study as it provides flexibility in array design, offers custom genotyping, and is appropriate for candidate gene association studies as in the present study. The details of different SNP genotyping methods are presented in **Appendix D**. The custom VeraCode GoldenGate

Assay on the BeadXpress platform which support mid- to low-plex (48 to 384-plex) genotyping panels was suitable for this study as in total 96 SNPs were genotyped.

The VeraCode technology is based on the generation of distinct digital holographic codes within each microbead, which ensures production of robust, precise and accurate data. It consists of cylindrical glass microbeads, 240 x 28 microns, each capable of carrying high density 24-bit codes to produce a large number of bead types. Upon laser excitation, each bead diffracts the beam into multiple components, generating a digital signature, which is captured by the camera in the BeadXpress Reader System, emitting a unique holographic code (**Figure 2.5**). The digital coding enables not only the identification of the target but also the sample ID, the kit as well as the laboratory ID. Desired multiplex levels ranging from one to several hundred are achieved by pooling uniquely coded non-overlapping beads, providing specific detection.

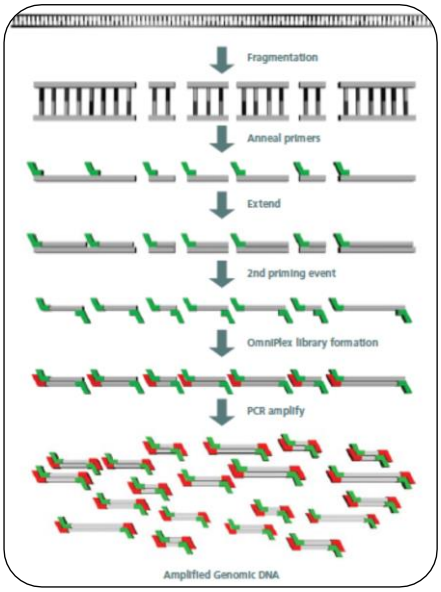


Figure 2.4 Overview of GenomePlex WGA Kit Process**.** The process starts with the fragmentation of the DNA into short overlapping strand templates followed by two priming steps of these short strands leading to Omniplex library generation. PCR amplification is performed using the DNA template in the library pool. Purification of the amplified genomic DNA is undertaken once the amplification is complete. (Source: Sigma Technical Notes 2007).

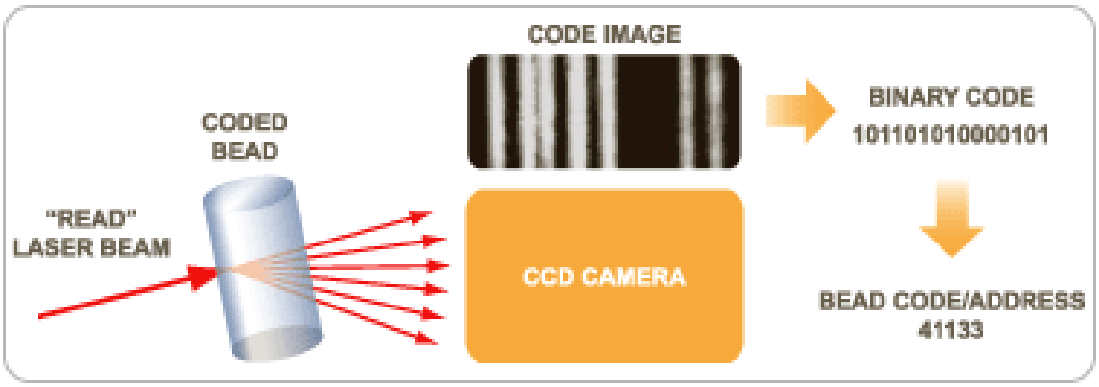


Figure 2.5 Schematic diagram showing generation of digital holographic code using VeraCode**.** Each bead contains a specific digital code, which diffracts the laser beam into various individual components. These components structure form the digital signature of the bead code when read by the CCD camera of the BeadXpress reader. (Source: Illumina Technical Note, 2007).

The BeadXpress reader evaluates the result generated through Veracode microbeads. It utilizes a two-fold colour identification system where it reads the distinct holographic code (code data), which is embedded in every Veracode microbead and also identifies the signal strength (fluorescence data) connected to the bead. Because the assay developed with VeraCode microbeads typically includes up to 30 replicates of each bead type, the data generated by the BeadXpress Reader is particularly robust. Each of the 30 replicates of the micrboead is optically visualized for a minimum of 12 times, generating up to 300 individual data set points for each sample. In addition, the decoding procedure also serves as a robust quality control measurement demonstrating the hybridization efficiency of all the beads in the assay (Fan et al. 2005).

2.8.8 SNPs Selection

The BeadXpress platform is preferred mainly due to its cost effectiveness, specifically for 96 SNP/sample tests. Therefore, based on their known or putative role, 66 SNPs from 16 candidate genes were selected for the SNP analysis based on previously published data as discussed in **Table 1.3** and **Table 1.4**. These genes can be subdivided into four broad groups: (1) genes primarily involved in xenobiotic

metabolism pathway; (2) antioxidants genes; (3) genes primarily involved in nitrosative stress; and 4) asthma related genes (**Table 2.2**). Additionally, a group of 30 random SNPs was included to provide information on population stratification. One marker from the long arm and the short arm of each autosome was selected using computer generated random numbers to identify the genomic position. Only SNPs not known to be associated with any specific disease and showing the greatest frequency variation across the HapMap populations (within 20kb of the selected position) were included in the ancestry informative marker SNP set. More information on the SNP selection criteria are included in Annex H.

These 96 SNPs were then examined for modification effects of air pollution exposure on respiratory health outcomes in the cross-sectional cohort London Low Emission Zone Schools study. Certain groups of SNPs on the array were selected a priori for use in the analysis of gene environment interactions on specific endpoints: For the lung function endpoints the modifying effects of SNPs in the glutathione transferases (GSTP1, GSTM1 and GSTT1) and NAD(P)H dehydrogenase, quinone 1 (NQO1), cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) and the aryl hydrocarbon receptor (AhR), were investigated; for urinary markers of oxidative stress, epoxide hydrolase 1, microsomal (EPHX1), nuclear factor (erythroid-derived 2)-like 2 (Nrf2), extra-cellular superoxide dismutase (SOD3) and glutamate-cysteine ligase catalytic/modifier subunit (GCLC and GCLM); and for FeNO, SNPs in arginase 1 (ARG1) and 2 (ARG2), as well as inducible nitric oxide synthase (NOS2). In addition, single SNPs in gasdermin A (GSDMA) and B (GSDMB) were related to the presence of childhood asthma and wheeze. In the work described in this thesis, I will only address the SNPs, related to the modification of respiratory symptoms, lung function and oxidative stress biomarkers of air pollution interactions. Further details on the selection criteria used to identify the SNPs and their linkage disequilibrium are presents in the later results chapters (3-5).

Table 2.2 List of Genes Genotyped in the Study

Group	Genes	Official Symbol (Other Symbol)	Chr	No. of SNPs
Phase I xenobiotic genes	Cytochrome P450, family 1, subfamily A, polypeptide 1	CYP1A1	15	5
Phase II xenobiotic genes	Glutathione S-transferase mu 1	GSTM1	1	1
Phase II xenobiotic genes	Glutathione S-transferase theta 1	GSTT1	22	1
Phase II xenobiotic genes	Glutathione S-transferase, pi 1	GSTP1	11	5
Phase II xenobiotic genes	NAD(P)H dehydrogenase, quinone 1	NQO1	16	4
Phase II xenobiotic genes	Epoxide hydrolase 1, microsomal (xenobiotic)	EPHX1	1	2
Antioxidant genes	Superoxide dismutase 3	SOD3 (EC-SOD)	4	6
Nitric oxide signalling	Nitric oxide synthase 2, inducible	NOS2 (iNOS)	17	14
Nitric oxide signalling	Arginase 1	ARG1	6	6
Nitric oxide signalling	Arginase 2	ARG2	14	9
Regulate phase I xenobiotic genes	Aryl hydrocarbon receptor	Ahr	7	5
Regulate phase I xenobiotic genes	Nuclear factor (erythroid-derived 2)-like 2	NFE2L2 (Nrf2)	2	3
Asthma related genes	Gasdermin A	GSDMA (GSDM1)	17	1
Asthma related genes	Gasdermin B	GSDMB (GSDML)	17	1
Glutathione-related metabolism	Glutamate-cysteine ligase catalytic subunit	GCLC	6	1
Glutathione-related metabolism	Glutamate-cysteine ligase modifier subunit	GCLM	1	2
Ancestry markers	Random SNPs		Various	30
Total number of SNPs studied				96

2.8.9 Oligo Pool Assay Design

Each custom genotyping kit contains Oligo Pool Assays (OPAs) assessed through Illumina's Assay Design Tool (ADT), to verify the functionality of the probes on GoldenGate technology, avoiding the necessity for multiplexed PCR optimization. Illumina allocates designability ranks and SNPs scores to demonstrate the ability to achieve successful assay. Initially, a total of 100 SNPs were selected for evaluation by Illumina. The final selection of SNPs for genotyping samples was based on a score of >0.6 and designability rank of 1.0, as these oligos were most likely to achieve success during the genotyping for whole genome amplified DNA (Cunningham et al. 2008). Although this is a quality parameter, it does not guarantee the success of the assay. Therefore, upon receipt of the sequence scoring, the SNPs that failed to achieve the minimum score (0.6) and designability rank less than 1.0 were excluded from the list and replaced by the ones which were initially short listed but not selected. SNP with failure reflected a number of issues including close proximity to another SNP of interest with a distance of just 61 nucleotides and a minor allele frequency (MAF) less than 0.05. As the success of the OPAs design depends on each SNP probe sequence as they form a mixture in one solution, this process was performed numerous times until a good match of sequences was achieved. Finally, 96 SNPs were selected using the following criteria: designability rank of 1.0 and a minimum MAF of at least 0.1 in one of the HapMap populations. Following this, a final Score file was produced for the requisition of order. A system was developed of assigning names to all SNPs from SNP1-SNP96, which included relevant data for quick identification of each SNP. **Appendix D** shows reference sequence (rs), location of the SNP in the genomic DNA and polymorphic site in bracket with forward slash separating the two alleles. Most of the SNPs are located at the intergenic region.

2.8.10 llumina GoldenGate Genotyping Assay

All the samples were genotyped at the Genomics Centre, Guy's Hospital using the protocol described previously (Shen et al. 2005). The assays were conducted in 96 well plates. Each protocol ran for 3 days and two plates were processed simultaneously. DNA samples were added to each of the wells and the previously designed OPAs were then added for hybridization of the oligos. As all SNPs possess three oligos therefore 96 OPAs contain 288 probes in the VeraCode solution. Two of the three probes are exclusively for the two potential alleles at the SNP site, Allele Specific Oligos (ASO), which bind at the SNP site and the third binds downstream to the SNP site, the Locus Specific Oligo (LSO). Every 3' end ASO sequence will hybridize to the target sequence that will match up to either of the two alleles, and the 5' end will bind to a universal PCR primer sequence. The LSO possesses three regions. The 5' end is used for hybridizing downstream of the SNP, the central one is the unique address sequence and the 3' end corresponds to the universal PCR primer sequence. The unique address sequence present on an individual bead in the Sentrix BeadChip, is later utilized for the identification of each genotyped SNP.

In the primer hybridization process, the oligos hybridized to the genomic DNA sample bound to paramagnetic beads. As hybridization took place prior to amplification, it removes amplification bias in the assay. Subsequent to hybridization, several washing steps were performed to reduce noise by removing excess and mis-hybridized oligos. The information about the genotype was obtained by extension of appropriate ASO and ligation of the extended full-length product to the LSO providing a template for PCR using universal PCR primers P1, P2 and P3. As the ligation depends on the hybridization of both oligos, this becomes a significantly important step. The primers P1 and P2 distinctly match up with one of the corresponding ASOs. Thus, just one primer will bind to the oligonucleotide depending on the hybridized ASO at the target DNA. In addition, primers P1 and P2 are labeled with Cy3- and Cy-5 in such a way that the dye ratio is suggestive of allele frequency. PCR amplification was then performed to enrich the number of labeled

SNPs available for the allele identification. The single strand DNA was hybridized to beads on the Sentrix Array Matrix through their unique address sequence. Following hybridization, the array was scanned using the BeadXpress reader to analyze the fluorescence signal. Each address corresponds to a specific locus, and the existence of either Cy3, Cy5, or both signals on a specified bead type representing either AA, BB or AB genotypes. Each genotype call was then assessed utilizing Illumina GenomeStudio software (**Figure 2.6**). In total, 988 samples were genotyped on twelve separate plates. On a final plate, 29 poorly performing samples were re-run to fill up the remaining wells.

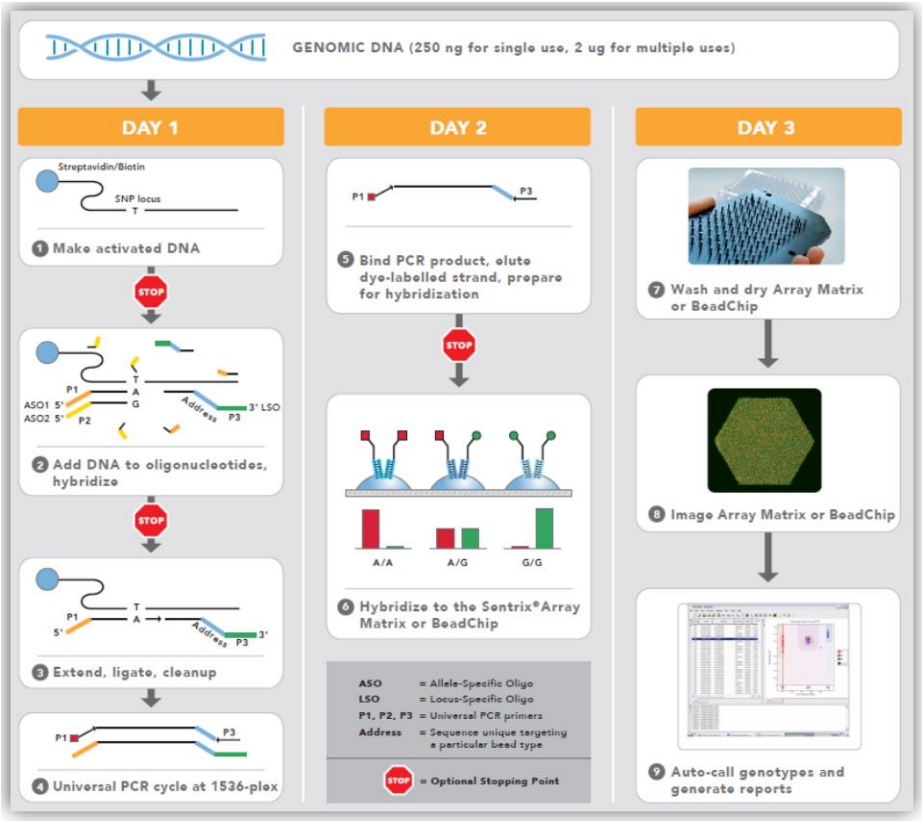


Figure 2.6 Illumina GoldenGate assay work flow. The overall assay takes place in approximately 9 steps over 3 days. Firstly, the DNA is activated which is then added to the allele specific (ASO) and locus specific (LSO) oligonucleotides and hybridized. Subsequently, allelic specific extension and ligation is performed to generate the DNA template for amplification process. During this step, fluorescence label is incorporated for the identification of specific allele for the particular SNP. Hybridization of the amplicon with the VeraCode beads is accomplished and scanned using BeadXpress Reader. Intensities data is analyzed using GenomeStudio Software for automated genotype calling and clustering. (Source: Illumina Technical Note 2006).

2.8.11 Genotype Calling

The intensity data obtained from each plate was assessed using the Genotyping module (v1.9.4) of the Illumina GenomeStudio (v2011.1) software. A preliminary run is performed by selecting ‘‘Cluster SNPs’’ by using GenCall cutoff of 0.25 as recommended for GoldenGate Assay. Following that raw data is generated through automated genotype calling.

The GenomeStudio clusters the samples as “call” or “no call” in a Polar plot depending on the zygosity of the loci of the SNP genotypes. Here “calls” represent loci having GenCall score above the call region threshold; in this case, 0.25 while “no calls” are assigned to loci having GenCall scores below the call region threshold. Genotyping success rate was determined to demonstrate the quality of the genotyping by using the call rate measurement. Therefore, the call rate for each sample was calculated as follows: calls/ (no calls + calls) x 100. The outcomes are represented in GenomeStudio as GenoPlots or SNP Graphs (see **Figure 2.7**) where each point corresponds to a single locus. The polar plot shows the normalized theta value (B allele frequency) on the X-axis (the angle of deviation from the pure A allele) corrected for the cluster position for the particular SNP of interest. The theta values range between 0 and 1 where value 0 refers to pure A signal and 1 refers to pure B signal. The Y-axis is based on normalized total intensity (R) for two alleles (A and B) for the given SNP. Therefore, Log R ratio ($R_{observed}/R_{expected}$) is the normalized R value divided by the expected normalized R value. GenCall scoring for diploid samples normally produces three clusters of alleles; each distinguished by a corresponding distinct colour coding where AA (red) and BB (blue) represent homozygous allele and AB (purple) represents heterozygous allele and the scores range from 0 to 1. Scores 0 and 1 correspond to AA and BB homozygous alleles respectively whereas samples lying in between these two regions with score 0.5 represent heterozygous AB allele. Samples falling anywhere in between these three clusters are referred to as missing data.

2.8.12 Quality Control Analysis

Sample-dependent, sample-independent, and contamination controls were included in the OPAs to assess the performance of the GoldenGate assay. The controls were viewed using the GenomeStudio software to examine the reliability of the assays, reagents, BeadChips and equipment, in addition to the samples. **Figure 2.8** shows typical images of controls from the GenomeStudio dashboard reporting information on allele-specific extension, PCR Uniformity, gender, extension gap, first hybridization, and second hybridization controls. The gender controls are used mainly to ascertain assay success and/or sample quality, such that failure to observe signal may indicate general process failure in any steps of the assay, or poor DNA sample quality. In addition, any gender mismatch of reported sex to the inferred sex based on the genotyping on the X-chromosome denotes possible sample contamination during the sampling or/and genotyping process. Since there are only two bead types for the gender control assays, the measurement is acknowledged to be prone to error. For this reason I did not a priori exclude gender mismatch samples from the analytical data set, as this does not infer that the other SNPs are in error. Rather, all analysis were ran with and without the 41 mismatched samples (<5% of all typed samples), to evaluate the extent that this influenced the associations observed. As no influence was observed here we report these data within the full analysis.

The use of various controls for testing efficiency of allele specific extension, PCR amplification efficiency, sample sex verification, gap extension efficiency, annealing specificity and hybridization specificity are important in the assessment of overall performance of the samples, reagents, equipment and BeadChips. These controls are already built into the GoldenGate assay (OPA) and is observed using GenomeStudio by different dashboards. Controls dashboard for the current analysis in presented in **Appendix E**.

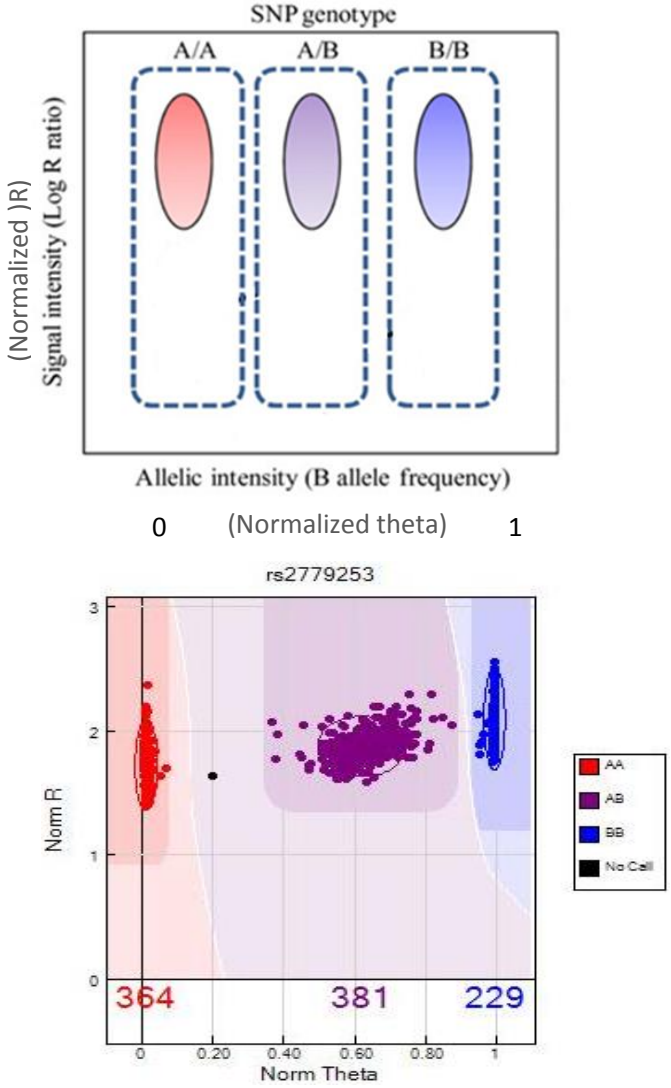


Figure 2.7 Genoplot. *Illustration of a GenoPlot (top panel).* Alleles are coded with specific colours, red for allele AA, purple for AB the heterozygous allele and blue for BB, the homozygous allele. Genotypes are determined using signal intensity (Norm R, Y-axis) and B Allele frequency (Norm Theta, X-axis) for particular SNP. *Genotype calling using GenomeStudio software (bottom panel).* A representative Genoplot produced by BeadStudio for SNP rs2779253. Each dot represents an individual sample. Samples in the red zone are AA allele; samples in the purple zone are AB; and samples in the blue zone are BB allele. Based on the figure, a call score of 0 indicates that 364 subjects are AA, a call score of 0.5 indicaes 381 subjects have BB homozygous allele and a call score of 0.5 demonstrating 229 subjects with AB heterozygous allele. There is one individual with black dot representing missing data and assigned "no call".

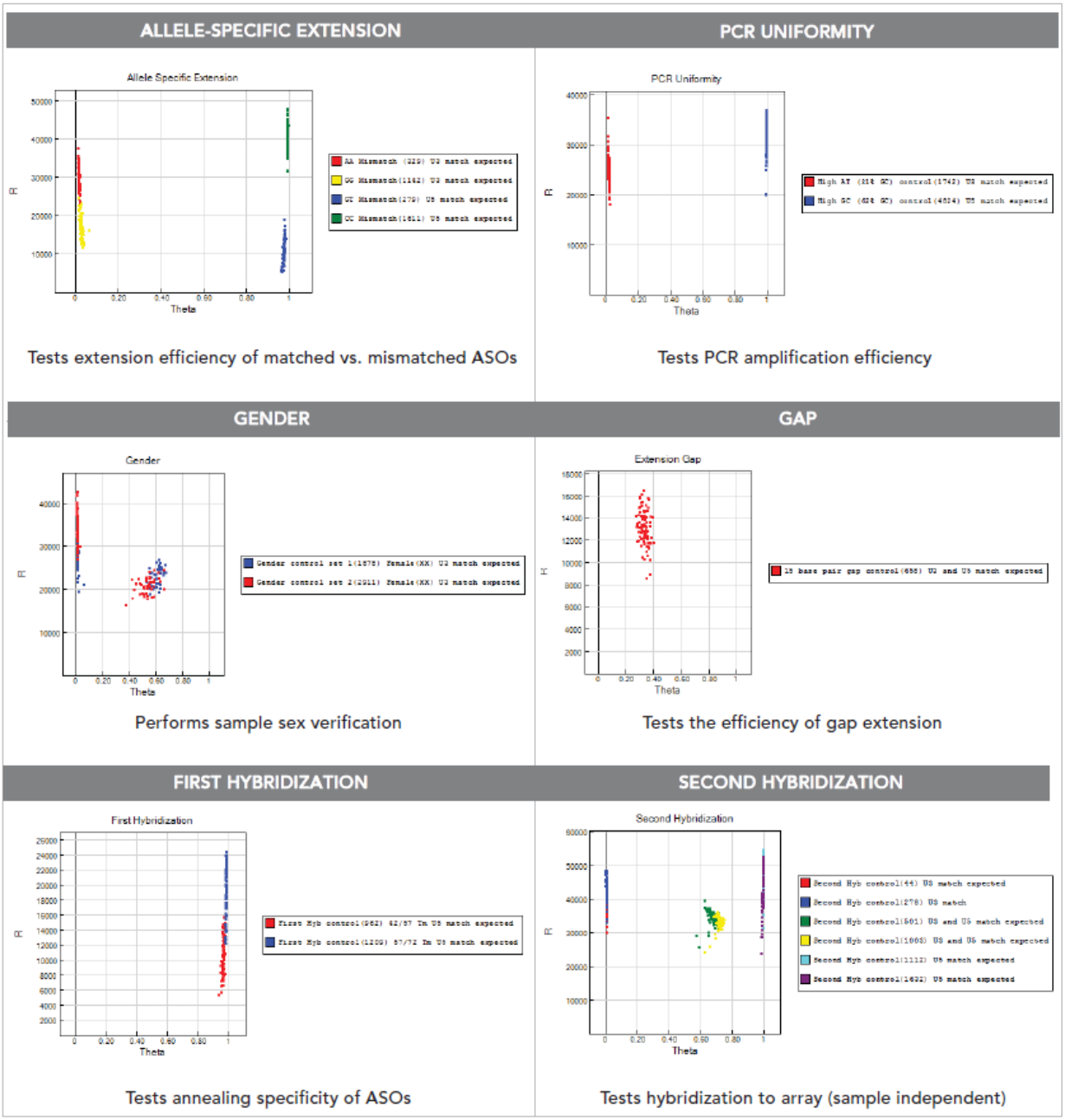


Figure 2.8 Typical Controls Dashboard Displayed in the GenomeStudio Module.

Despite a manufacturing issue in 2 plates (plate 5 and plate 11) where OPAs for controls were not added, most of the samples (>95%) showed good performance throughout the assay procedures and appeared to be within required specification, ensuring robustness of the assay to produce reliable genotype data (see **Appendix E** control dashboards for all samples in the study). Subsequently, all SNPs underwent manual investigation and adjustment (if needed) based on the following three criteria: (i) sample call rate; (ii) SNP call frequency and (iii) Hardy–Weinberg equilibrium. Sample call rates needed to achieve >0.90 with a GC cutoff score of 0.25, whereas SNPs call frequency should be >0.95 having cluster separation of approximately 0.25 or an average GenTrain score >0.60. In addition, Hardy–Weinberg equilibrium $\chi 2$ also needs to have p-values > 0.05 (detail description as in **Section 2.8.15**).

2.8.13 Evaluation of Samples

Investigation of samples that required reprocessing or removal from the analysis was performed based on the GenCall quality metric score, calculated for every genotype data point based on angle, dispersion, overlap of clusters and intensity. The samples with score equal to 1 are located in the center of the cluster, whilst a decreased GenCall score is indicative of a poorly performing sample. Scatter plots between sample call rate, as a function of sample number (see **Figure F1** & **F2** in **Appendix F**) and a 10% GenCall score as a function of the sample call rate (see **Figure F3** & **F4** in **Appendix F**) were generated to identify poorly performing samples. The problematic samples (possessing low sample call rate, low 10% GenCall scores, and outliers) were then reprocessed. In total 29 samples were reprocessed. Subsequently, the reprocessed samples were included for reevaluation of the sample performance and quality. Samples with call rate below 0.90 after reclustering were excluded from final data set.

2.8.14 Evaluation and Manual Editing of SNP Clusters

Manual editing is required with GoldenGate projects as not all SNPs perform well. Therefore, for the overall improvement in the call rates, it is important to detect the loci, which require manual editing or needs to be removed. This is generally achieved by evaluation of the reclustered SNPs based on the quality of the cluster separation. The quality of SNP clustering for a specific locus is determined by the GenTrain score (between 0 and 1). To begin with, the SNPs with the poorest clustering based on the GenTrain score are highlighted. SNPs possessing GenTrain scores of less than 0.6 were removed (see **Figure G1** in **Appendix G1**). Details on cluster separation, SNP frequency, GenTrain score, AB T mean, AB R mean and other distribution scores for all 96 SNPs are shown in **Table G1** in **Appendix G**. Each SNP showing low cluster separation was then investigated for cluster overlapping. The cluster separation determines the distance between the three genotype clusters (AA, BB and AB) based on the theta value ranging from 0 to 1 as explained in the **Section 2.5.4**. The clusters showing no overlapping were then manually revised to obtain a better separation, while the overlapping clusters were removed from the final dataset (see **Figure G2** in **Appendix G**). However, this indicator is far from ideal because cluster separation scores are derived from the degree of separation of the two homozygous clusters versus the heterozygous cluster instead of separation between the two of the homozygous clusters (Butler and Ragoussis 2008). Nevertheless, this method still yields useful evidence for the SNP quality and as shown in the current work that the SNPs having the cluster separation of even less than 0.25 may still act as a functional SNP.

SNPs call frequency is also an important parameter when assessing the quality of the SNPs. It is determined as the percentage of all samples at each locus having a call score above the 0.25 threshold. Therefore, the SNPs showing a low call frequency were also zeroed and removed from the analysis. Some of the SNPs in the present study also displayed a high heterozygous ratio (see **Figure G3** in **Appendix G**), which prevented manual re-clustering of these SNPs and they were therefore removed from the final analysis. Another criterion to evaluate the SNP functionality was the investigation of their

AB R and AB T values. A SNP with an AB R mean value (low intensities) too low to be effectively called were also removed from the analysis. On the other hand, evaluation of AB T values helped in identifying SNPs with a possible shift of the heterozygote cluster to the homozygote. For this purpose, SNPs having AB T values from 0 to 0.2 and 0.8 to 1, i.e. not clustered around the centre were analysed. For clusters that were difficult to edit, the locus was zeroed and removed from the final data (see **Figure G4** in **Appendix G**).

Using the internal controls is not mandatory in the genotyping workflow. However, by including a technical repeat plate it was possible to test the reproducibility and concordance of the data and confirm the robustness of the assay. This test was only performed following the manual editing and removal of poor quality samples and SNPs. Overall, technical repeat on both plates shows 100% reproducibility except for three samples. Two samples showed concordance of 0.99 whereas in another paired sample, the discordant SNP is simply because it was not called instead of miscalled.

In general, once the clusters were manually edited (many samples had call rates of 1.00) only 13 had call rates below 0.9 (of these, 3 had failed completely) and this is likely due to the samples themselves. Using the criteria outlined above, ten SNPs were classified as having failed: rs17072738 (SNP8: random SNP), rs1871042 (SNP28: GSTP1), rs2472299 (SNP68: CYP1A1), rs2074175 (SNP75: random SNP), rs8192290 (SNP53: EC-SOD), rs1138272 (SNP18: GSTP1), rs2234922 (SNP47: EPHX1), rs17883901 (SNP19: GCLC), rs6599689 (SNP94: random SNP) and rs4147581 (SNP4: GSTP1). The dataset significantly improved after these SNPs were zeroed, indicating poor performance of the individual SNPs prior to the manual editing. **Figure G5** & **Figure G6** in **Appendix G** shows 8 SNP graphs representing some of the successfully genotyped SNPs prior to editing and subsequent to manual editing.

Subsequently, using GenomeStudio, genders were estimated based on the intensities of X chromosome and the proportion of heterozygous genotype for this SNP. One would expect that samples marked as male have average intensities (R), half of the mean

intensities of female samples. This is because males are hemizygous at X chromosome SNPs. Generally, most of the males and females in the current project are grouped into two discrete clusters (see **Figure E4** in **Appendix E**). Overall, samples reported as "female" exhibit X chromosome and have intensity higher than the samples marked as "male". However, 41 (<5%) samples out of 974 appeared to be mismatched resulting in 22 males possessing X chromosome clustering at theta ~0 whilst 19 females showing the Y chromosome with theta value ~0.6. Consequently, all these mismatch samples were highlighted for sensitivity analysis. A test on equality of proportions using Stata 10.1 (StataCorp, College Station, TX, USA) was performed to compare genotype distribution for each SNP before and after exclusions of the gender-mismatched samples to examine if there would be any significance difference in the proportions. As expected, none of the genotype distributions was significant and the lowest p value was 0.521.

Finally, downstream analysis application requires investigation of the distinct base for each SNP of every sample. Therefore, the genotype data was transferred to Excel for the creation of the Final Report. Of the remaining 975 samples, one appeared to be a duplicate (HF039b) as it had been identified previously for labeling error during the Year 3 sample collection and subsequently removed from the list. It is extremely difficult if not impossible to include the genotype data for each subject from the population investigation. Therefore, the genotype data of the entire study population (more than 80,000 genotype data) is presented in the accompanying CD-ROM with the thesis.

2.8.15 Hardy Weinberg Equilibrium

Deviation from the Hardy Weinberg Equilibrium (HWE) may result from genotyping errors, inconsistency in the HWE assumptions, or simply by a random chance. The HWE frequency for a locus can be calculated by adding up the frequencies of all individual alleles on that locus. For example, considering p as the frequency of an allele

(A), and q the frequency for the alternative allele (a) then the expected HWE frequency for a locus containing two alleles of either genotypes AA, Aa or aa would be p^2, 2pq and q^2 respectively. For the HW equilibrium, the combined value of all the three individual genotypes should be equal to 1 as do the allele frequencies ($p^2 + 2pq + q^2 = 1$). Therefore, the measurement of deviation from the HWE is essential for quality control purpose.

The goodness-of-fit χ^2 test is the commonly used statistical tool to investigate the HWE based on the observed genotype frequencies from the dataset and expected genotype frequencies calculated from the HW equation. The null hypothesis is that alleles selection are randomly, resulting in expected HWE frequencies (p^2, 2pq and q^2) in the population. The alternative hypothesis is that the alleles selection are non-randomly leading to HW disequilibrium in the population. The HWE is calculated by χ^2 test using the following equation: n(n-1)/2df, with n being the number of allele at the locus, which is being investigated. Additionally, calculation of the degree of freedom (df) can be done by using the following equation: g-k, where g represents the number of genotypes while k stands for the number of alleles on each locus (Wittke-Thompson et al. 2005). For instance, the χ^2 test, which investigates the extent of deviation from the HWE, would provide df=1 for the SNP which consists of 2 alleles and 3 expected genotypes. In the current work, the statistical program STATA was utilized to conduct the standard χ^2 test on all of the SNP markers in the population to identify any departure from the HWE. The p-values <0.05 obtained from the χ^2 test would mean significant deviations from the HWE (Wigginton et al. 2005).

A significant finding during the QA/QC procedure was that approximately 28% (33 out of 86 SNPs) of the SNPs deviated from the HWE. Most of the association studies utilize repeat genotyping of some of the individuals as a quality control measure to identify the genotyping errors (Tintle et al. 2009). In the current study, the concordance rate for the repeat genotyping, (>99.99%) indicates that the deviation from the HWE is unlikely due to the genotyping errors. Since the study subjects were drawn from an ethnically diverse population, violations of the HWE may be related to the population stratification.

Therefore, HWE test was performed stratified by ethnicity. The stratification was performed based on the self-reported ethnicity where 4 major ethnicity groups were identified as: Asian, White, Black and Mixed as described in the **Table H1** in **Appendix H.** Detail on Hardy Weinberg calculation for each ethnicity is tabulated in **Table 2.3**. This approach decreased the number of SNPs deviating from HWE to 15 in total, with the Asian population having the largest number of SNPs not in the HWE. Further investigation revealed that eight out of ten SNPs deviated from the HWE NOS2 genotypes. Six of them (rs4795080, rs2779253, rs1889022, rs10853181, rs2531866, rs1014025, and rs2531872) were selected based on the SNPs being in high linkage disequilibrium in the haplotype block (Islam et al. 2010). The Asian group is known for non-random mating with either first cousin or close relatives. Although non-random mating does not change the allele frequencies, it could lead to alteration in the genotype frequencies resulting in violation of Hardy Weinberg assumptions. Whereas, 8 SNPs are deviated from HWE in Black ethnicity group with half of them are random SNPs. In White population, 3 SNPs out of 7 SNPs which deviated from HWE are also random SNPs. Meanwhile, in mixed ethnicity group, 5 SNPs do not conform to HWE.

In addition, another factor which may contribute to the deviation of Hardy Weinberg could be non-random patterns of missing data or systematic errors in genotyping (for an example inconsistent missing data in heterozygotes may cause an imbalance between heterozygous and homozygous leading to an excess of homozygous) (Wittke-Thompson et al. 2005). However, this effect is more prominent in case-control studies. Nevertheless, in this study, the exclusion of samples with genotyping call rate less than 0.90 was minimal, involving just 1.4% of the population study. Therefore, it is unlikely that missing data has altered the genotype frequencies for the particular SNPs.

A flow diagram outlining all of the QA/QC processes employed on the genetic data set and described in the preceding section is presented in **Figure 2.9**.

Table 2.3 Hardy Weinberg Equilibrium p-values of investigated SNPs

SNP Index	SNPs ID	Gene	Asian p-value	Black p-value	White p-value	Others p-value
SNP1	rs2917666	NQO1	0.331	0.693	0.476	0.453
SNP2	rs7782389	Random SNP	0.053	0.076	0.163	0.227
SNP3	rs997279	Random SNP	0.450	0.020	0.356	0.318
SNP5	rs3759757	ARG2	0.119	0.153	1.000	0.136
SNP6	rs1137933	iNOS (NOS2)	0.282	0.256	0.359	1.000
SNP7	rs2826003	Random SNP	0.627	0.518	0.749	0.811
SNP9	rs2246012	ARG1	0.436	0.095	0.358	0.460
SNP10	rs2074113	Ahr	0.709	1.000	0.719	0.432
SNP11	rs7232792	Random SNP	0.327	0.069	0.028	1.000
SNP12	rs2284659	EC-SOD (SOD3)	0.389	0.001	0.519	0.569
SNP13	rs2781668	ARG1	0.334	0.044	0.815	0.689
SNP14	rs2066853	Ahr	0.583	0.300	0.350	0.802
SNP15	rs2733262	Random SNP	0.802	0.208	0.324	1.000
SNP17	rs17599586	ARG1	0.645	1.000	0.590	0.735
SNP18	rs742869	ARG2	0.637	1.000	0.703	0.046
SNP20	rs2364723	Nrf2	0.064	0.438	0.880	0.839
SNP21	rs8078340	iNOS (NOS2)	0.039	0.368	0.543	1.000
SNP22	rs3742879	ARG2	1.000	0.222	0.547	1.000
SNP23	rs268691	Random SNP	0.278	0.087	0.713	0.573
SNP24	rs2301022	GCLM	0.133	0.524	0.183	0.162
SNP25	rs3170633	GCLM	0.375	0.448	1.000	0.178
SNP26	rs2779251	iNOS (NOS2)	0.005	0.587	0.245	1.000
SNP27	rs7132743	Random SNP	1.000	0.012	1.000	0.660
SNP29	rs6770096	Random SNP	0.228	0.858	1.000	1.000
SNP30	rs7557529	Nrf2	0.495	1.000	0.328	0.048
SNP31	rs2779253	iNOS (NOS2)	0.001	0.866	0.779	0.354
SNP32	rs2606345	CYP1A1	0.463	0.142	1.000	0.578
SNP33	rs8192287	EC-SOD (SOD3)	0.196	0.897	0.048	1.000
SNP34	rs2781666	ARG1	0.538	0.754	0.263	0.186

Note: SNPs which are not in HWE ($p<0.05$) are highlighted in orange.

Table 2.3 Hardy Weinberg Equilibrium p-values of investigated SNPs

SNP Index	SNP ID	Gene	Asian p-value	Black p-value	White p-value	Others p-value
SNP35	rs4795080	iNOS (NOS2)	0.003	0.328	0.388	1.000
SNP36	rs7144243	ARG2	1.000	0.854	0.571	0.059
SNP37	rs17722841	Ahr	0.600	0.696	0.403	0.562
SNP38	rs699473	EC-SOD (SOD3)	0.520	0.736	0.513	0.572
SNP39	rs1889022	iNOS (NOS2)	0.001	0.865	0.576	0.354
SNP40	rs9282799	iNOS (NOS2)	0.651	0.364	0.975	0.887
SNP41	rs7591449	Random SNP	0.651	0.606	0.445	0.074
SNP42	rs3894194	GSDM1	0.450	0.443	0.535	0.036
SNP43	rs1442293	Random SNP	0.763	0.202	1.000	0.239
SNP44	rs749174	GSTP1	0.278	0.863	0.791	1.000
SNP45	rs689453	NQO1	0.906	0.528	0.554	0.519
SNP46	rs3796644	Random SNP	0.906	0.846	1.000	0.052
SNP48	rs17779352	Ahr	0.264	0.575	0.705	0.659
SNP49	rs2128238	Random SNP	0.406	0.137	0.217	0.690
SNP50	rs1799814	CYP1A1	0.065	0.821	0.023	0.267
SNP51	rs10454231	Random SNP	0.762	0.020	0.353	1.000
SNP52	rs7776785	Random SNP	0.219	0.363	0.421	0.304
SNP54	rs7140310	ARG2	0.431	0.186	0.603	0.189
SNP55	rs1800566	NQO1	0.905	0.644	0.837	0.345
SNP56	rs17861115	CYP1A1	0.133	0.444	0.364	0.051
SNP57	rs11735827	Random SNP	1.000	0.389	0.461	0.432
SNP58	rs2531866	iNOS (NOS2)	0.001	0.768	0.779	0.346
SNP59	rs9900426	Random SNP	0.198	1.000	0.895	0.026
SNP60	rs10517	NQO1	0.405	1.000	0.190	1.000
SNP61	rs2282885	Ahr	1.000	0.582	0.896	0.124
SNP62	rs2001350	Nrf2	0.614	0.776	0.335	0.254
SNP63	rs527705	Random SNP	0.666	0.049	0.390	0.705
SNP64	rs4902503	ARG2	0.122	0.400	1.000	0.194
SNP65	rs7875663	Random SNP	0.743	0.210	0.324	0.239

Note: SNPs which are not in HWE ($p<0.05$) are highlighted in orange.

Table 2.3 Hardy Weinberg Equilibrium p-values of investigated SNPs

SNP Index	SNP ID	Gene	Asian p-value	Black p-value	White p-value	Others p-value
SNP66	rs407257	GSTT1	0.000	0.000	0.000	0.000
SNP67	rs2531872	iNOS (NOS2)	0.001	0.470	0.781	0.254
SNP69	rs221454	Random SNP	0.517	0.420	0.620	0.706
SNP70	rs2781667	ARG1	0.538	0.877	0.322	0.181
SNP71	rs13306703	EC-SOD (SOD3)	1.000	0.730	0.000	0.300
SNP72	rs2198843	CYP1A1	0.489	1.000	0.546	0.137
SNP73	rs778233	Random SNP	0.583	0.193	0.902	1.000
SNP74	rs7156352	ARG2	0.489	1.000	0.151	0.118
SNP76	rs3824781	Random SNP	0.645	0.230	0.414	0.420
SNP77	rs2236687	Random SNP	0.281	0.621	0.902	0.708
SNP78	rs10459953	iNOS (NOS2)	0.111	0.747	1.000	0.829
SNP79	rs10483801	ARG2	0.021	1.000	0.401	1.000
SNP80	rs2153747	Random SNP	0.563	0.180	0.043	1.000
SNP81	rs10853181	iNOS (NOS2)	0.001	0.769	0.674	0.346
SNP82	rs1014025	iNOS (NOS2)	0.228	0.208	0.220	0.845
SNP83	rs2779248	iNOS (NOS2)	0.159	0.700	1.000	0.695
SNP84	rs8192288	EC-SOD (SOD3)	0.196	0.898	0.056	1.000
SNP85	rs470411	Random SNP	0.293	0.757	0.417	0.823
SNP86	rs9394047	Random SNP	0.767	0.530	0.278	0.689
SNP87	rs11809289	Random SNP	0.280	0.268	0.044	0.432
SNP88	rs12885261	ARG2	0.517	0.701	0.380	0.451
SNP89	rs2297518	iNOS (NOS2)	0.675	0.546	0.130	1.000
SNP90	rs1695	GSTP1	0.889	0.123	0.595	1.000
SNP91	rs2305480	GSDMB	0.610	0.375	0.804	1.000
SNP92	rs2781659	ARG1	0.535	0.786	0.256	0.442
SNP93	rs366631	GSTM1	-	-	-	-
SNP95	rs1051740	EPHX1	0.364	0.043	0.480	0.429
SNP96	rs6017870	Random SNP	0.874	0.237	0.085	1.000

Note: SNPs which are not in HWE ($p<0.05$) are highlighted in orange.

Moreover, genetic variation, caused by an increase or decrease in the allele frequencies, may also cause deviation from HWE. Furthermore, gene flow, representative

of transfer of new alleles to a population, caused by mating between two populations could also potentially cause HW disequilibrium. This may explain deviation of HW in 'others' group in five SNPs (rs742869, rs7557529, rs3894194, rs8192290 and rs407257) as this population mainly consists of mixed ethnicity and we are unable to stratify them to any specific ethnicity.

The GSTM1 variant is a copy number variant. The null genotype has no copies of the GSMT1 gene and therefore reduced expression/functional activity for this enzyme. The alternate alleles may have 1, 2 or more copies. The pseudo-SNP rs366631 which represents a biomarker of GSTM1 deletion was selected for this study based on Huang et al. (2009) and only distinguishes whether it is homozygous deletion (null) or not. This pseudo-SNP is therefore unable to differentiate between the numbers of alternative alleles. It does not therefore behave like a true SNP, which is usually bi-allelic with 3 genotype options. The reason for using this SNP was the existence of the complex duplication/deletion patterns in the GSTM1 gene. There are 2 regions, one upstream and one within GSTM1 gene with shared high homology. Since there are common deletions in the GSTM1 gene, when both copies of the upstream region were deleted, the genotype call will be made solely based on homozygous variants (TT). When one copy of upstream region is deleted, the heterozygous genotype call (CT) is made (personal communication with Huang). By using the similar GoldenGate technology which was applied in the Huang et al. (2009) study, the genotype call is made for this pseudo-SNP to infer the GSTM1 deletion. Consequently, as this variant is multi-allelic, it does not necessarily have to conform to Hardy-Weinberg principle, which refers to truly bi-allelic variants.

Nonetheless, there is no simple single explanation for the observed deviation from the HWE. The major limitation of the HW investigation is that it fails to provide the basis of the deviation in the data and the only important information obtained from the test is whether or not the data meet the expected proportions calculated for HW. This observation may also occur by chance with one in 20 SNPs deviating from the HWE at the 5% level.

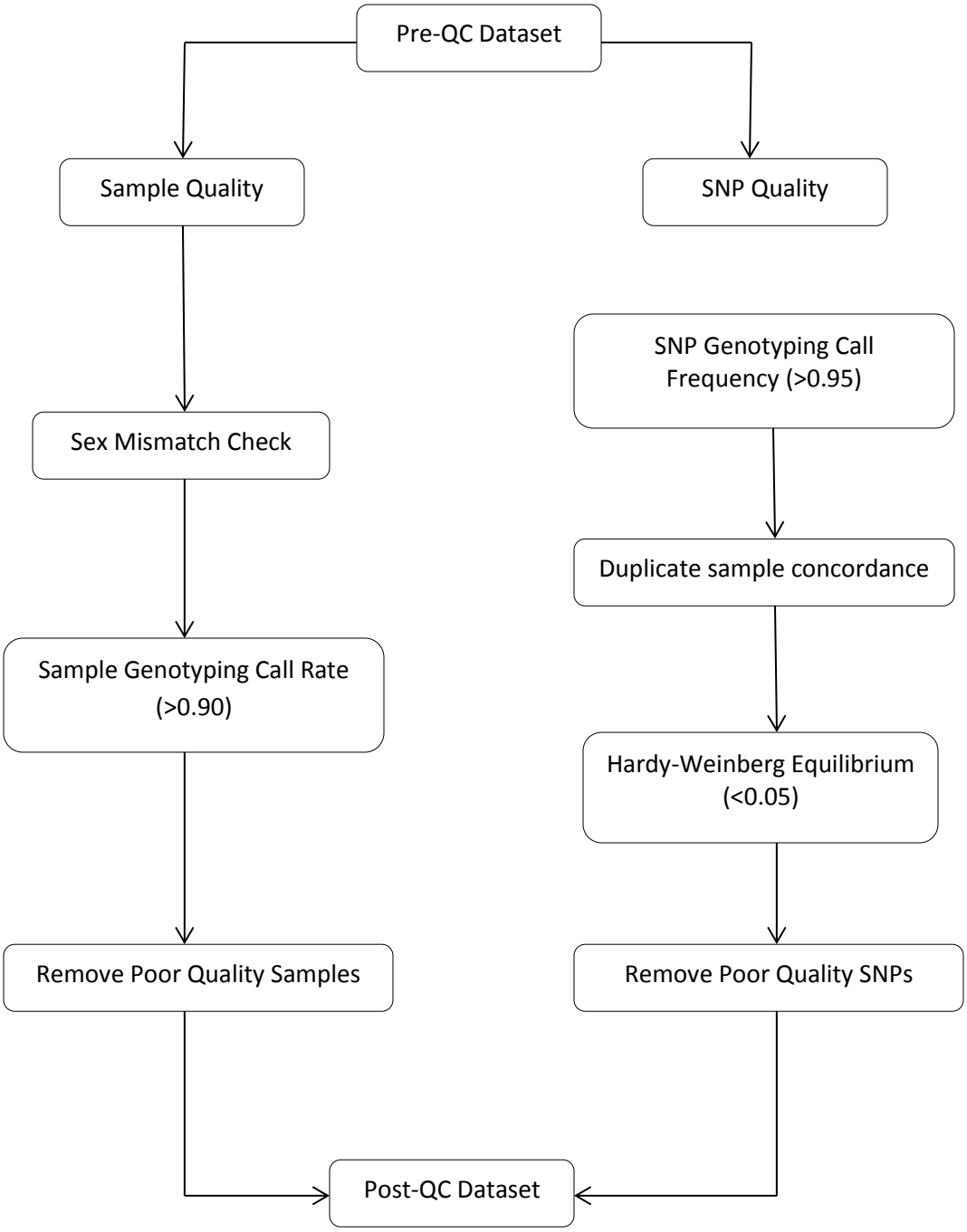


Figure 2.9 A flow chart overview of the QA/QC process for the genetic dataset.

2.9 Population Structure Analysis

As the study population consisted of multi- and mixed ethnic subjects, therefore, an attempt was made to categorize the study population to 'true' subpopulations based on genetic variations of ancestry markers. Population structure was investigated by employing a Bayesian clustering algorithm in the STRUCTURE software using genotype data derived from a set of random SNPs. Genotype data from random SNPs in all 974 subjects were analyzed to assign individuals into probable population clusters. However, assignment of the individuals to two or more populations occurs when their genotype data are consistent with admixture. In this population stratification exercise, analysis was based on an 11 subpopulations classification (**Table H1** in **Appendix H**), based on self-reported ethnicity. This model is based on the assumption that there are a unknown number (K) of population clusters in each data set defined by allele frequencies at each locus reflective of population ethnicity/ continental origin (Pritchard et al. 2000).

The admixture model was implemented in the STRUCTURE 2.3.3 to run a 100,000 iteration (20,000 burn in) using Bayes' Rule Markov Chain Monte Carlo (MCMC) model with correlated allele frequencies and clustering without a priori information on sampling location. The main presumption of the admixture model was that individuals from unrelated populations might share ancestry that enables identification of subpopulation clusters in the study population. Therefore, by using the MCMC method and investigating the range of LnPr(D) a more accurate K value based on the highest likelihood at optimal K was identified. MCMC was run using the same iterations and burn in period at several values of K (Pritchard et al. 2007). This procedure was repeated using the same iterations and burn in period at least 10 times for each cluster with the K values between K=2 and K=5 yielding the most probable number of populations. This step is also important to validate the pattern of the population structure generated. In STRUCTURE, log likelihood, InP(D) referring to the probability Pr(X|K) given the data (X) forms the basis of the K value estimation. Visual examination of the plot of the log probability was performed to determine the most

appropriate K value as an ad hoc indicator and plateau method based on the Falush et al. (2007) was used to confirm the convergence. Subsequently, determination of K value was targeted to the smallest K value "that captures the major structure in the data" (Pritchard et al. 2007) within the plateau region. In addition, the highest likelihood value was considered the most likely of the K value for the population investigated. Details on the analysis are presented in **Appendix H**.

Chapter 3

Air Pollution, the London Low Emission Zone, and the Prevalence of Respiratory and Allergic Symptoms in Schoolchildren in East London

3.1 INTRODUCTION

The International Study of Asthma and Allergies in Childhood (ISAAC) reported that Great Britain had the highest prevalence of asthma symptoms in children, worldwide (The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee 1998). According to the most recent available data from the National Health Service in England, in 2011-12, 5.9% of patients or 3.3 million people had asthma and in 2010-11, there were 23,452 emergency admissions to hospital of children with asthma (NHS Information Centre for Health and Social Care 2011). The recent Health Effects Institute review of the evidence relating traffic-related air pollution to health concluded that there is sufficient evidence to link traffic with exacerbation of asthma symptoms in children, and suggestive evidence for a link to asthma onset in childhood (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010).

Levels of traffic-related air pollution in London are the worst in the UK and among the worst in Europe, with European Union (EU) limit values for particulate matter with an aerodynamic diameter of less than 10μm (PM_{10}) and nitrogen dioxide (NO_2) being exceeded in many areas of the city. Recent prospective cohort studies in the United States have demonstrated that chronic exposure to traffic emissions has significant negative impacts on lung development during childhood, resulting in clinically important deficits

into adolescence (Gauderman et al. 2004; Gauderman et al. 2007). A recent study of children in the UK also demonstrated impaired lung function in children with a higher density of particulate carbon in their airway macrophages (Kulkarni et al. 2006). There are also significant short-term effects of air pollution on health, notably the exacerbation of symptoms in asthma sufferers. A number of studies have shown associations between increased concentrations of traffic-related air pollutants and hospital admissions for asthma in children (Delfino et al. 2009; Iskandar et al. 2012; Lee et al. 2006; Samoli et al. 2011). A study of adults with mild-to-moderate asthma demonstrated a significant reduction in lung function after walking on a busy London street for two hours, associated with exposures to ultrafine particles and elemental carbon – both pollutants arising from diesel exhaust (McCreanor et al. 2007).

In February 2008, a low emission zone (LEZ) covering most of Greater London was established, with the aim of reducing traffic-related air pollution and thereby improving air quality and the health of Londoners. The London LEZ is the largest in the world, covering an area of 2,644 km^2 with a population of 8.2 million (ONS 2012). Modelling work undertaken to assess the potential impact of the LEZ (comparing scenarios with and without it) predicted reductions in emissions of PM_{10} from road traffic sources in Greater London of 2% in 2008 and 6% in 2012, with 4% and 10% reductions in emissions of nitrogen oxides (NO_x) from road traffic in 2010 and 2012, following implementation of phases 1-3 of the scheme (Transport for London 2008). Greater reductions were predicted in areas where concentrations of PM_{10} and/or nitrogen dioxide (NO_2) exceeded EU limit values prior to the scheme's implementation. The LEZ study was therefore initiated to assess the accrued health benefits of living within the LEZ on the respiratory health of 8-9 year old schoolchildren in the east London boroughs of Hackney and Tower Hamlets. These boroughs are also areas of significant socioeconomic deprivation (Department of Health 2012a, b), with 77% and 67% of the populations in Hackney and Tower Hamlets, respectively, living in the most deprived areas, based on national quintiles of the 2010 Index of Multiple deprivation in England (Department for Communities and Local

Government 2011). As our study was only initiated after phase 1 and 2 of the scheme, in November 2008 and due to the delay in the implementation of phase 3 from October 2010 to January 2012, the first three years of the life time of the scheme are treated here as a baseline study examining the associations between air pollution and respiratory/allergic symptoms prior to the interventions in January 2012.

In addition, we examined the associations between characteristics that could be described as an asthma 'genotype' (single nucleotide polymorphisms, SNPs, in gasdermin genes (gasdermin A, 17q21.1 and gasdermin B, 17q12), previously shown to be associated with child-onset asthma (Moffatt et al. 2010)) or 'phenotype' (significant bronchodilator response (Dundas et al. 2005)) and the prevalence of wheeze and asthma. It was logistically impossible to include a comparable control group in a low air pollution area.

We found that current rhinitis symptoms were associated with traffic-related air pollutants, that asthma 'genotype' was associated with lifetime asthma, and asthma 'phenotype' was associated with lifetime asthma and current wheeze symptoms. We did not observe any year-on-year improvements in air pollution concentrations over the first three years of London's LEZ or any reduction in the prevalence of respiratory/allergic symptoms.

3.2 METHODS

3.2.1 Study Design and Recruitment

Schools in Tower Hamlets and Hackney were selected to achieve a range of annual exposures to NO_2 (based on dispersion models (Transport for London 2008)). These schools were approached and asked to participate; in those, which agreed, all Year 4 children (aged 8-9) were invited to participate. This was a sequential cross-sectional study, with data collected over three consecutive winters (Nov-Mar, 2008-11). Information about the study was sent home with each child, along with a consent form and a questionnaire for parents to

complete and return to the school. During a single study visit to each school each year, health assessments were conducted to examine lung function and collect biological samples, and completed questionnaires were collected. Parents were required to give their written consent, and children their verbal assent, to participate in the health assessment. The return of a completed questionnaire was taken to imply consent to use the information provided therein. The study was approved by the local research ethics committee (East London & The City HA Local Research Ethics Committee 2, REC Ref Number 08-H0704-139) and conformed to the standards set by the Declaration of Helsinki.

3.2.2 Questionnaire

The questionnaire included three sections; the core questions on asthma, rhinitis, and eczema from the ISAAC questionnaire for 6-7 year olds (The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee 1998), see **Panel A**. Information on the questionnaires was entered exactly as recorded, regardless of apparent inconsistencies and whether or not a question should have been answered according to the instructions within the questionnaire. Where a question was not answered, it was coded as 'missing'. Details of missing and inconsistent data for each section of the questionnaire are reported in the **Appendix I**. Symptom prevalences were calculated by dividing the number of positive responses to each question by the total number of completed questionnaires. The definitions used for current symptoms are outlined in **Table 3.1**.

Table 3.1 Definitions of current and lifetime respiratory/allergic symptoms

Symptom	Definition
Current	
Wheeze	Has your child had wheezing or whistling in the chest in the past 12 months?
Severe wheeze	Has your child had wheezing or whistling in the chest in the last 12 months? AND AT LEAST ONE OF: ≥4 attacks of wheezing in the past 12 months; ≥1night/week of sleep disturbed by wheezing in the past 12 months; wheezing severe enough to limit speech to only one or two words at a time between breaths in the past 12 months.
Rhinitis	In the past 12 months, has your child had a problem with sneezing, or a runny, or a blocked nose when he/she DID NOT have a cold or the flu?
Eczema	Has your child ever had an itchy rash which was coming and going for at least 6 months? AND Has your child had this itchy rash at any time in the last 12 months? AND Has this itchy rash at any time affected any of the following places: the folds of the elbows, behind the knees, in front of the ankles, under the buttocks, or around the neck, ears or eyes?
Lifetime	
Asthma	Has your child ever had asthma?
Hayfever	Has your child ever had hay fever?
Eczema	Has your child ever had eczema?

3.2.3 Health Assessment

3.2.3.1 Demographics

Health assessments were conducted during the study visit to each school. Date of birth, home address and ethnicity were obtained from school records. Age was calculated as the date of the study visit to the school minus the date of birth. A deprivation score was assigned to the home postcode using the 2010 Index of Multiple Deprivation. Height in cm was measured using a Stadiometer and weight in kg was measured using a standard weighing scales. Body mass index (BMI) was calculated as body mass (kg) divided by height (m) squared.

Panel A: Questions provided in the study packs for the children's parents/guardian

Core questions for asthma (from ISAAC questionnaire for 6-7 year-olds)

1. Has your child ever had wheezing or whistling in the chest at any time in the past? Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6
2. Has your child had wheezing or whistling in the chest in the past 12 months? Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6
3. How many attacks of wheezing has your child had in the past 12 months? None/1-3/4-12/More than 12
4. In the past 12 months, how often, on average, has your child's sleep been disturbed due to wheezing? Never woken with wheezing/Less than one night per week/One or more nights per week
5. In the past 12 months, has wheezing ever been severe enough to limit your child's speech to only one or two words at a time between breaths? Yes/No
6. Has your child ever had asthma? Yes/No
7. In the past 12 months, has your child's chest sounded wheezy during or after exercise? Yes/No
8. In the past 12 months, has your child had a dry cough at night, apart from a cough associated with a cold or chest infection? Yes/No

Core questions for rhinitis (from ISAAC questionnaire for 6-7 year-olds)

1. Has your child ever had a problem with sneezing, or a runny nose, or blocked nose when he/she DID NOT have a cold or the flu? Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6
2. In the past 12 months, has your child had a problem with sneezing, or a runny nose, or blocked nose when he/she DID NOT have a cold or the flu? Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6
3. In the past 12 months, has this nose problem been accompanied by itchy-watery eyes? Yes/No
4. In which of the past 12 months did this nose problem occur? (Please tick any which apply) January/February/March/April/May/June/July/August/September/October/November/ December
5. In the past 12 months, how much did this nose problem interfere with your child's daily activities? Not at all/A little/A moderate amount/A lot
6. Has your child ever had hayfever? Yes/No

Core questions for eczema (from ISAAC questionnaire for 6-7 year-olds)

1. Has your child ever had an itchy rash which was coming and going for at least six months? Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 7
2. Has your child had this itchy rash at any time in the past 12 months?Yes/No
   IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 7
3. Has this itchy rash at any time affected any of the following places: the folds of the elbows, behind the knees, in front of the ankles, under the buttocks, or around the neck, ears or eyes? Yes/No
4. At what age did this itchy rash first occur? Under 2 years/Age 2-4 years/Age 5 or more
5. Has this rash cleared completely at any time during the past 12 months? Yes/No
6. In the past 12 months, how often, on average, has your child been kept awake at night by this itchy rash? Never in the past 12 months/Less than one night per week/One or more nights per week
7. Has your child ever had eczema? Yes/No

3.2.3.2 Urine Assays

Spot urine samples were collected and placed on ice for the duration of the visit, then frozen and stored at -80°C on return to the laboratory. Urine cotinine concentrations were determined using a commercial microplate enzyme immunoassay (EIA) kit (Cozart Forensic Microplate EIA for cotinine, product no. M155B1) from Concateno (Abingdon, UK). Urine was analysed undiluted, as per the kit instructions. This kit is designed to be semi-quantitative, i.e. to give a positive or negative result, based on a cut-off point of 25 ng/mL (sample is positive if value > 25 ng/mL). A correction for individual variation in urine concentration was made using the cotinine to creatinine ratio (CCR) in ng/mg (calculated as: 100 x cotinine in ng/ml divided by creatinine in mg/dL), using a cut-off of value of 30 ng/mg to indicate a positive result (Henderson et al. 1989; Ehrlich et al. 1992). Urinary creatinine concentrations were also determined using a commercially available kit based on the Jaffe reaction (Cayman Chemical Company, Ann Arbor, MI, USA). Measurements were repeated if duplicates were either side of kit cut-off value (e.g. 24 and 26 ng/mL), or if duplicates were on one side of cut-off but SD was large. 34 samples out of 914 (3.7%) should have been repeated according to these criteria, but repeat samples were not available for 6 samples. All urinary analyses were performed by Dr Helen Wood at King's College London.

3.2.3.3 DNA Collection and Genotyping Assay

Saliva were collected using the Oragene DNA kit OG-250 (DNA Genotek Inc, Canada) as a source of DNA for genotyping. The collection was carried out according to the manufacturer's instructions. Samples were placed on ice for the duration of the visit, then frozen and stored at -80°C on return to the laboratory. The collection was carried out according to the manufacturer's instruction. Subsequently, genomic DNA was recovered from the saliva using salt out method, with genomic DNA concentrations and quality assessed using Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies,

Wilmington, DE) and Quant-iT™ PicoGreen® assay (Invitrogen). In addition, integrity and quality of DNA was confirmed by gel electrophoresis. Whilst the genomic DNA obtained was of a good quality and integrity, the yield requirement for the subsequent SNP analysis required whole genome amplification on all collected samples by using the GenomePlex Complete Whole Genome Amplification (WGA2) Kit (Sigma Aldrich) as previously described (Arneson et al. 2008). Polymorphisms (SNPs) in gasdermin A (GSDMA) rs3894194 Arg18Gln (C/T) and gasdermin B (GSDMB) rs2305480 Pro298Ser (C/T) were analyzed by GoldenGate genotyping assay on IlluminaBeadXpress platform (Illumina Inc., San Diego, USA) and the genotype data was analyzed for quality control using the BeadStudio software.

3.2.3.4 Lung Function

Children's respiratory function was assessed by spirometry (Microlab, Micromedical, Carefusion), performed by trained investigators according to ATS-ERS guidelines (2005) with baseline and post-bronchodilator measurements, following salbutamol 400 μg administered by large volume spacer. Before each measurement volume calibration with a 3L syringe was undertaken. A sterile disposable filter/ mouthpiece was attached to the spirometer for each child and the equipment wiped with alcohol wipes between subjects. A maximum of 10 attempts was normally made (exceptionally, more if needed) until three acceptable and two repeatable attempts were attained. Each spirometry assessment aimed to obtain three acceptable and two repeatable attempts both pre and post bronchodilator. A short acting bronchodilator (salbutamol) was administered after baseline spirometry. Four 100 microgram actuations were given from a metered dose inhaler (MDI) through a spacer device (Volumatic). The children took 4 tidal breaths through the Volumatic spacer after each actuation. A minimum of 15 minutes later post bronchodilator spirometry took place. Quality control was based on the ATS/ERS guidelines (2005), modified for children (Praud 2004). Additional acceptability requirements were: rapid onset

of expiration, high well defined peak flow and a clear plateau on volume-time curve together with no evidence of cough, glottis closure or leak during the maneuver (from Asthma UK/ Growing Lungs Guidelines). For quality control and reporting purposes spirometry results were uploaded from the 3 study spirometers to a Carefusion program "Spirometry" where individual inspection of efforts was undertaken in detail. Reporting of results was according to ATS-ERS Guidelines and the best overall individual effort from each child both pre and post salbutamol was selected as the highest value of FEV_1 and FVC, together with FEF 25-75 from the best combined effort, reported from technically acceptable data, which met reproducibility requirements of 0.15 L agreement for FEV_1 and FVC for 2 efforts. Data were then extracted into Excel and individual efforts were cross checked with original data before being exported into the main study Access database. Raw data were transformed into Z (or standard deviation) scores using most the "All-Age" equations by Stanojevic et al. (2009) available for White subjects aged 3 to 80 years of age.

3.2.4 Air Pollution Modelling

Annual NOx, NO_2, PM_{10} and $PM_{2.5}$ concentrations were derived using the KCL urban model (Kelly et al. 2011; Beevers et al. 2013) using ADMS dispersion model v4 and road source model v2.3 (CERC19), measured hourly meteorological data, empirically derived NO-NO_2-O_3 and PM relationships and emissions from the London Atmospheric Emissions Inventory (GLA 2008). Separate annual surface maps for Greater London were produced for the years 2008, 2009 and 2010, available at a 20 m×20-m grid point resolution, linked to the following health assessment periods: Nov 2008 – March 2009, Nov 2009 – April 2010 and Nov 2010 – April 2011 respectively. Each yearly model reflected a range of pollutant sources and emissions, including major and minor roads, with detailed information on vehicle stock, traffic flows, and speed on a link-by-link basis. Other sources within the model included large and small regulated industrial processes, boiler plants, domestic and commercial combustion sources, agriculture, rail, ships, airports, and

pollution carried into the area by prevailing winds. A comprehensive description of this model has been published previously, together with information on validation against measurements (Beevers et al. 2013) and its performance against other urban dispersion models (Carslaw 2011). All exposures were based on the annual mean within a 20m radius buffer zone around the residential address of the child. Initial investigations using the year 1 data set examined the use of wider buffer zones, 50 and 100m found a high degree of correlation between the exposure estimates and therefore only the 20m buffer zone was employed in the final analysis (**Figures 3.1 – 3.4**). In addition, exposure estimates were weighted for periods spent at the home (H) and school (S) address points based on the following criteria: that each child spent 84.4% of their time at home and 15.6% at school, based on a 7 hour school day, for 5 days per week, 39 weeks per year. Thus each child's weighted exposure was estimated by $E_{H+S} = 0.884*E_H + 0.156*E_S$. Again, as these exposure estimates were highly correlated with the home exposure attributions only the later were employed in the analysis of the final data set (**Appendix I, Figures I1-I4**).

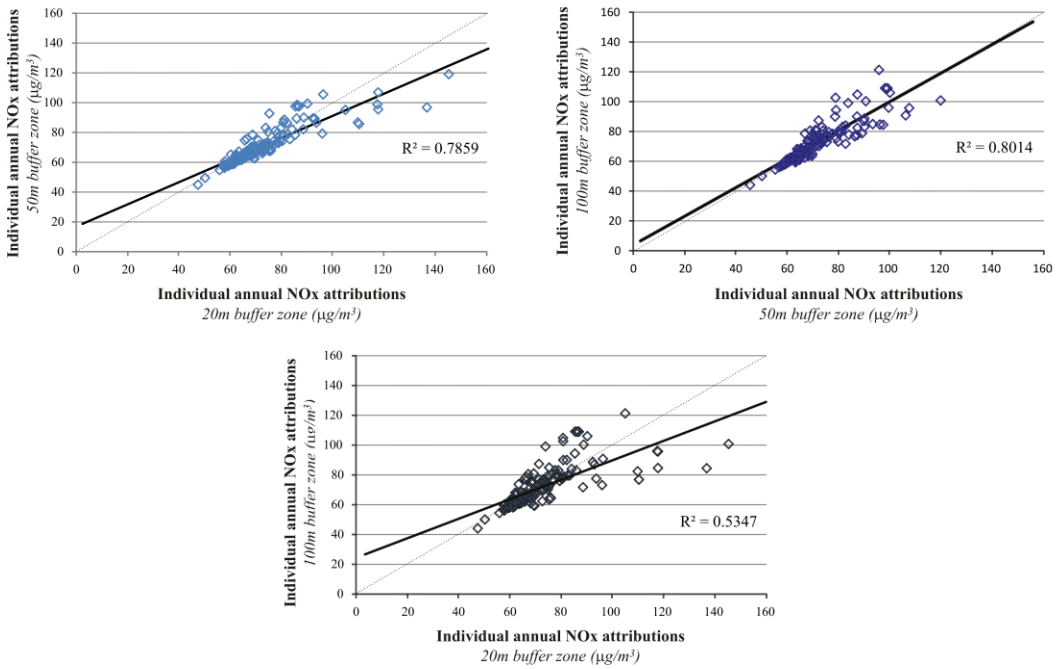


Figure 3.1 Correlations between annual NOx concentrations for the subjects recruited into the first year of the study, n=196, based on various buffer radii around their residential address. Data are presented with their linear trend line, together with the line of identify (hatched line).

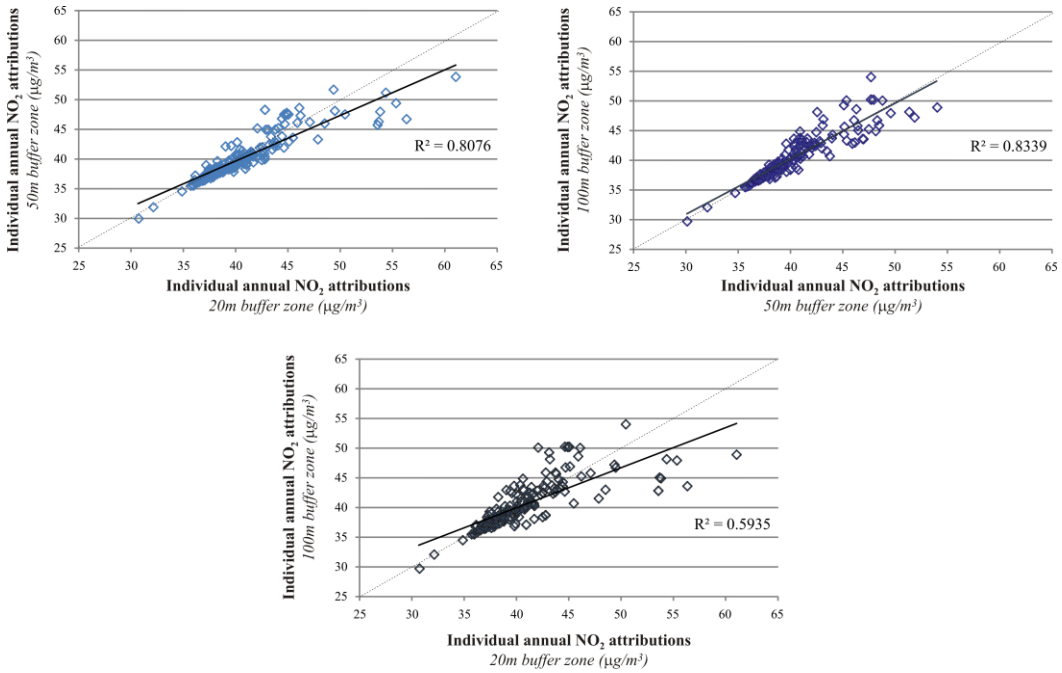


Figure 3.2 Correlations between annual NO_2 concentrations for the subjects recruited into the first year of the study, n=196, based on various buffer radii around their residential address.

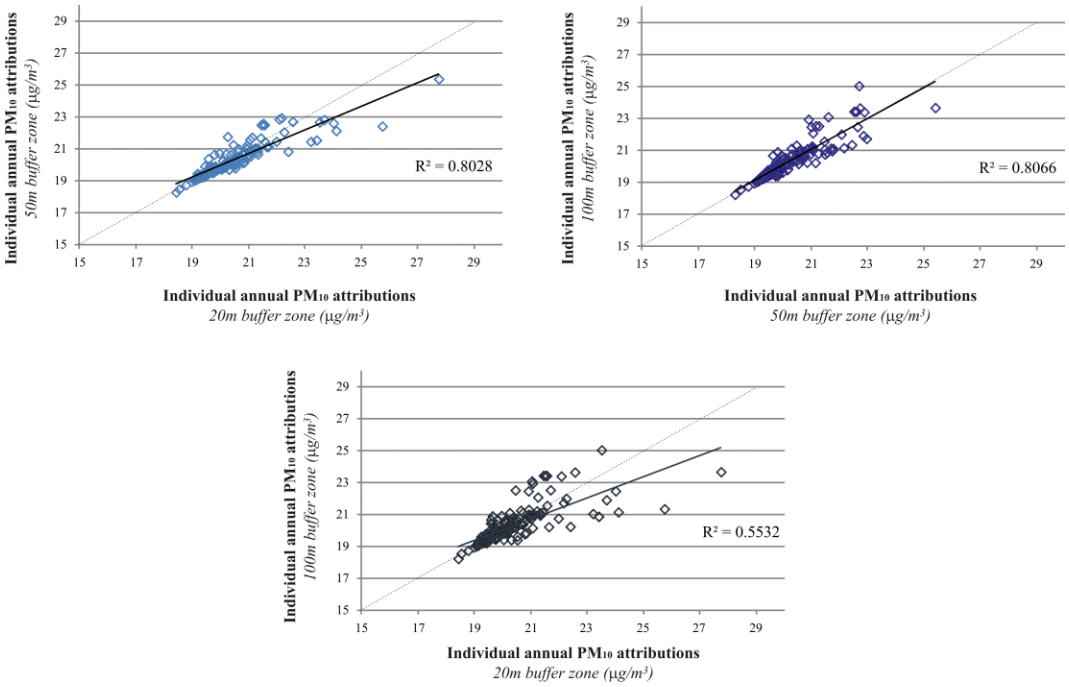


Figure 3.3 Correlations between annual PM_{10} concentrations for the subjects recruited into the first year of the study, n=196, based on various buffer radii around their residential address.

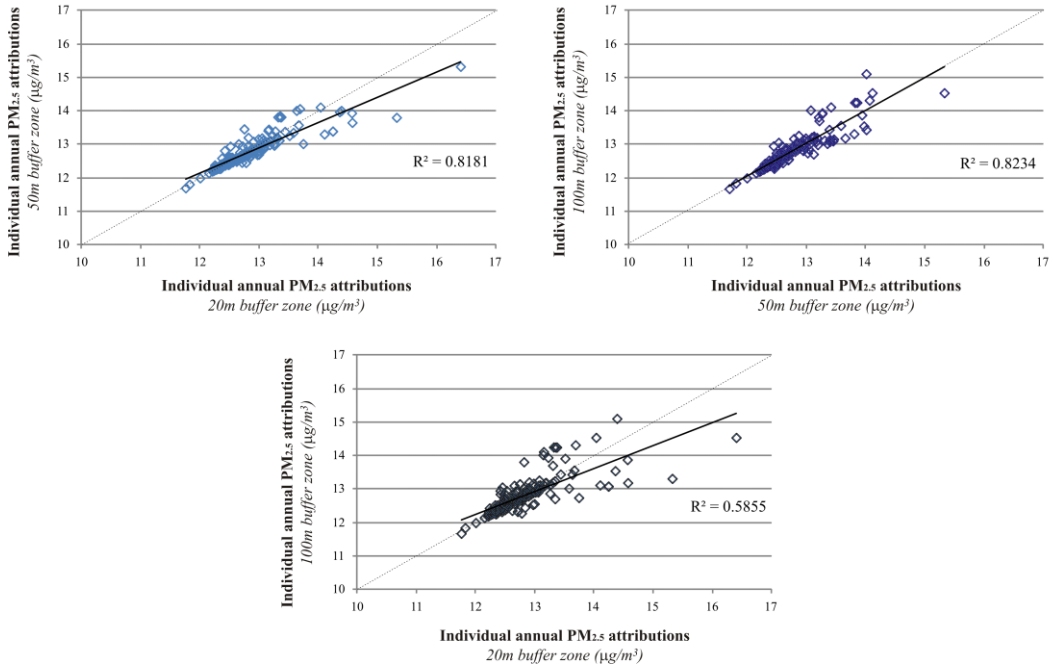


Figure 3.4 Correlations between annual $PM_{2.5}$ concentrations for the subjects recruited into the first year of the study, n=196, based on various buffer radii around their residential address.

3.2.5 Statistical Models

The primary outcome of the study was to detect a 4% improvement year-on-year in the average forced expiratory volume in 1 sec (FEV_1), hence the sample size was determined according to this variable rather than respiratory/allergic symptom prevalence. It was initially determined that 200 children would be needed per year, but this was revised upwards following the first year of data collection. From Year 1 data, the mean FEV_1 was 1.71 litres (SD=0.276 litres) and valid measurements of FEV_1 were obtained from 150/202 i.e. 74% of children. In order to detect a 4% increase in mean FEV_1, from 1.71 to 1.79 litres in two successive years, with 80% power at the 5% significance level, would require 245 children under simple random sampling. Assuming 74% (from Year 1) of children give valid FEV_1 measurements; the cluster size would be 22 children (out of a class of 30). The design effect, based on an intra-cluster correlation coefficient (ICC) of 0.02, is 1 + (22-1)*0.03 = 1.63, inflating the sample size to 399 children per year. Therefore, the number of classes to be sampled in Year 2 and subsequently is 399/22 = 19 classes (rounded up from 18.1).

All respiratory symptoms analysed were recorded as binary variables. Therefore, a multilevel mixed effects logistic regression model was applied allowing a random effect for school. Wherever enough outcomes were observed, we added age, sex, BMI, IMD score, ethnicity and ETS exposure to each model. If the observed outcome frequency was too low, we added only a subset of the covariates to the model to avoid over-fitting. To account for any potential year-on-year changes we also included study year as a coefficient in each model and individual measurements of air pollutant exposure were added to the models where applicable. The effects of asthma 'genotype' and 'phenotype' on the prevalence of current wheeze and lifetime asthma were assessed in those individuals reporting symptoms only, using separate models including genotype for GSDMA and GSDMB or bronchodilator response (BDR, expressed both as a continuous variable and as a yes/no variable with a positive value being ≥9% (Dundas et al. 2005)), excluding air pollution

exposures. PM_{10} was subsequently included in the model to examine any interactions between asthma 'genotype' and 'phenotype' and air pollution exposure. All analysis was performed using Stata 10.1. Models were examined at the 5% significance level for two-sided tests.

3.3 RESULTS

3.3.1 Demographics

Completed questionnaires were returned by 995 Year 4 pupils from a total of 1808 at the 23 participating schools (55% response rate). Of those 995, 901 were returned by pupils who also completed the health assessment and 94 questionnaires were returned by pupils who either declined to participate in the health assessment, or were absent from school on the day of the study visit (**Figure 3.5**). The participants are described in **Table 3.2**. Information on sex and ethnicity was missing for a small number of children (1.7 and 1.4%, respectively).

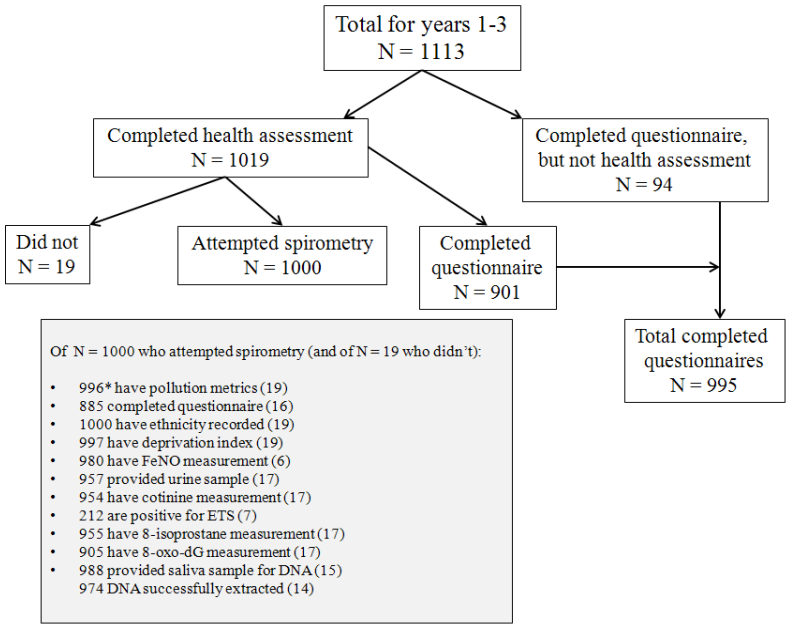


Figure 3.5 Data and sample availability for the subjects recruited into the study across years 1-3. The inset box provides detailed information on the 1000 subjects who attempted spirometry.

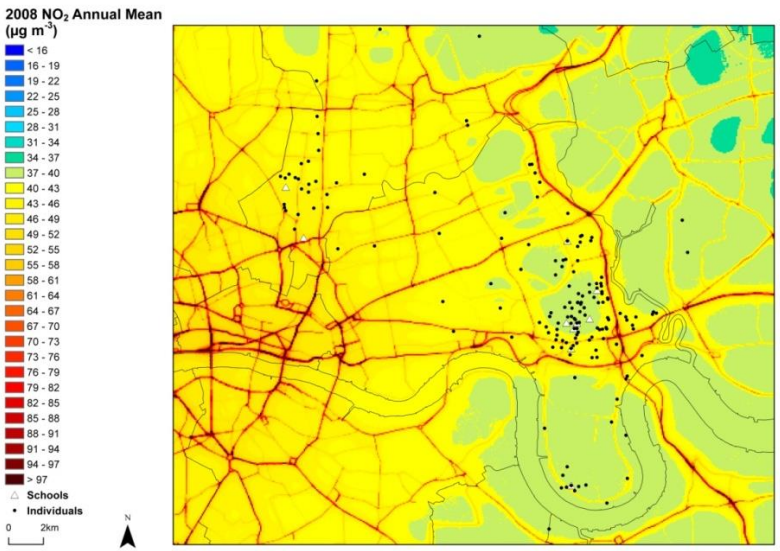
Table 3.2 Characteristics of the study sample, for all respondents and for those with current symptoms (within the last 12 months)

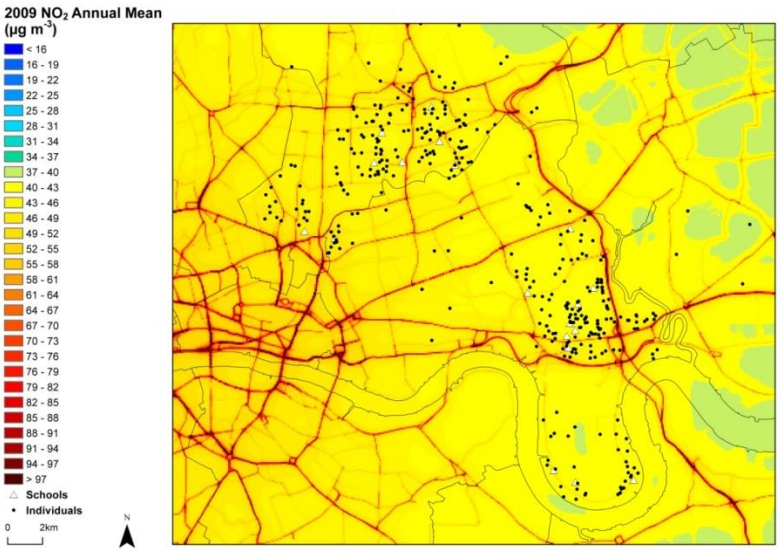
	Total	Wheeze	Rhinoconjunctivitis	Eczema
All (%)	995 (100)	111 (11.2)	125 (12.6)	148 (14.9)
Sex				
Male (%)	497 (49.9)	70 (14.1)	70 (14.1)	69 (13.9)
Female (%)	481 (48.3)	40 (8.3)	54 (11.2)	76 (15.8)
Not specified (%)	17 (1.7)	1 (5.9)	1 (5.9)	3 (17.6)
Ethnicity				
Asian (%)	360 (36.2)	34 (9.4)	48 (38.4)	57 (15.8)
Black (%)	240 (24.1)	24 (10.0)	30 (24.0)	45 (18.8)
White (%)	268 (26.9)	41 (15.3)	33 (26.4)	31 (11.6)
Mixed/other (%)	113 (11.4)	12 (10.6)	13 (10.4)	12 (10.6)
Not specified (%)	14 (1.4)	0	1 (0.8)	3 (21.4)

Percentages are for rows, except for first column, which reads vertically (i.e. percent of all respondents); percentages may not add to 100.0 due to rounding.

3.3.2 Modeled Exposures

The annual dispersion models employed for the exposure attributions are illustrated in **Figures 3.6 – 3.9**, for NOx, NO_2, PM_{10} and $PM_{2.5}$, with the address point for each subject and school illustrated. These surface maps of the study area clearly demonstrate the marked yearly difference in modelled concentrations, with the air quality in 2009, being significantly worse than that in 2008 and 2010. With reference to the NO_2 map for 2009 (**Figure 3.6**) it was significant that all but one of the subjects lived at locations that failed the attain the UK limit value for annual NO_2, 40 $\mu g/m^3$. Across the three years of the study over 85% of the study volunteers lived in areas non-compliant with the UK limit value. The distribution of annual pollutant concentrations (over years 1-3) is illustrated in **Figure 3.10**, with summary statistics.





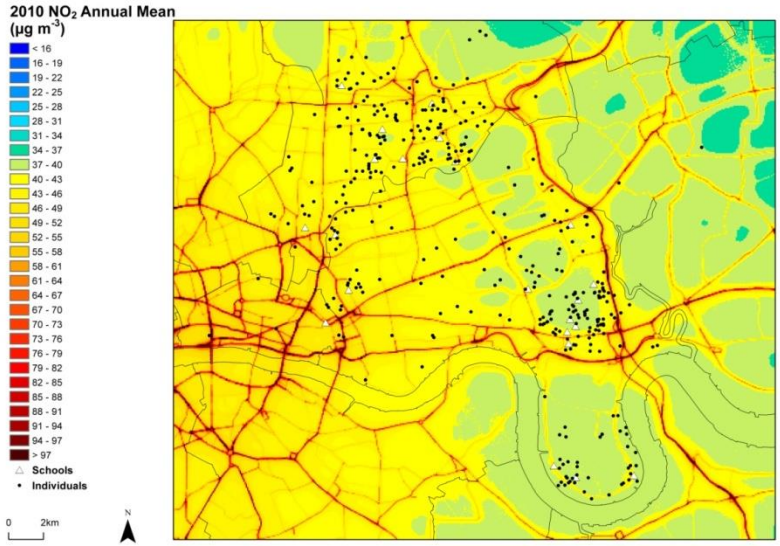


Figure 3.6 Annual NO_2 maps covering the London boroughs of Hackney and Tower Hamlets for the years 2008 -2010. Individual points reflect the residential address of the volunteers tested during the winter periods, with the school locations highlighted using open triangles.

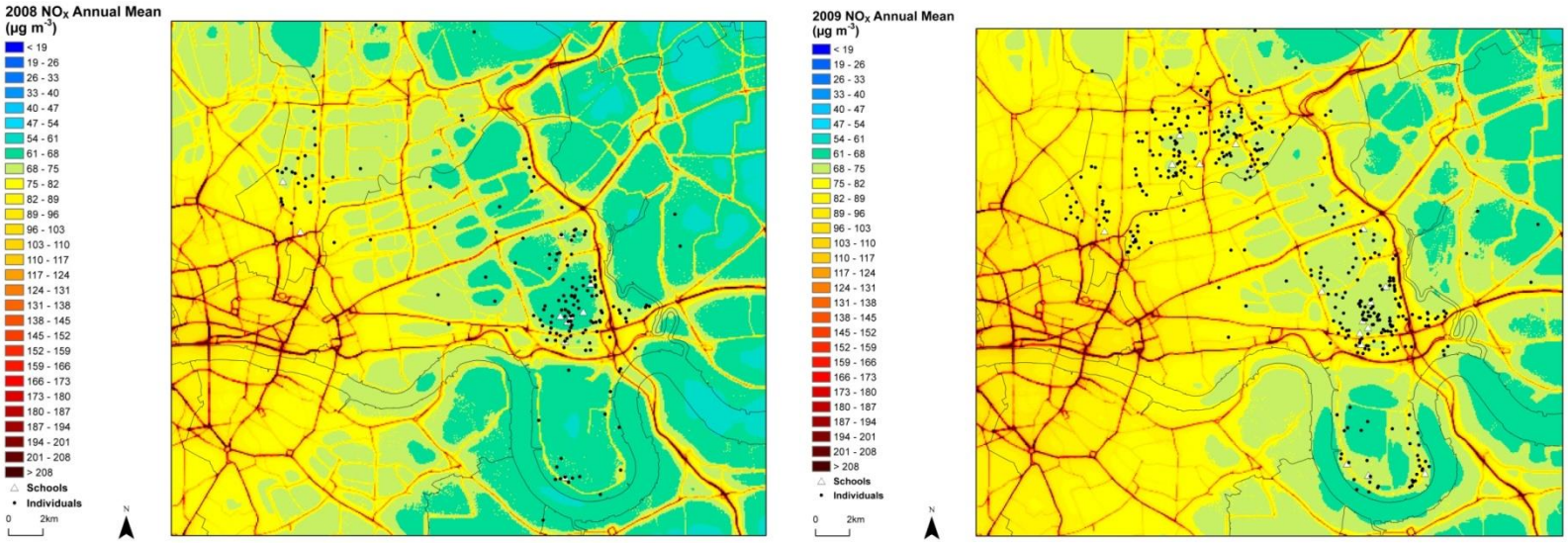
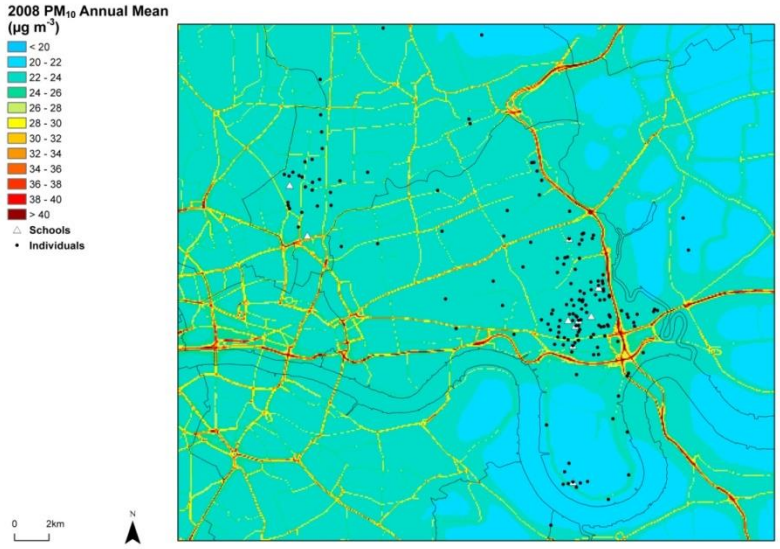
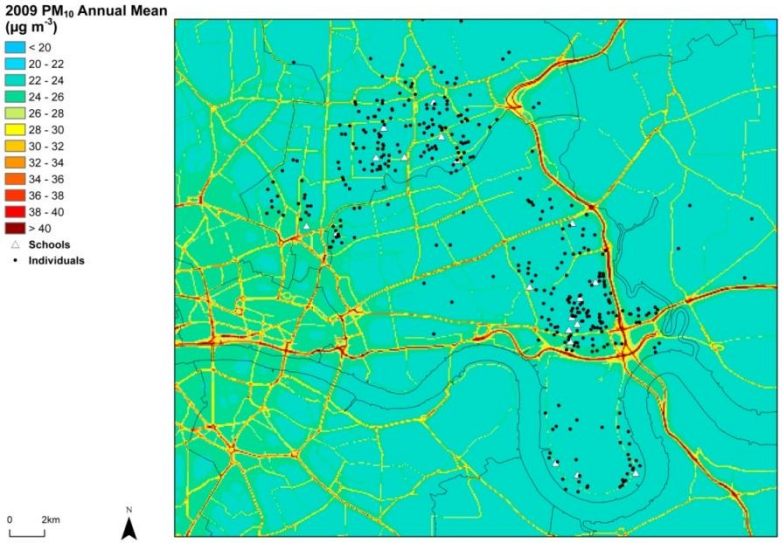


Figure 3.7 Annual NOx maps covering the London boroughs of Hackney and Tower Hamlets for the years 2008 -2010. Individual points reflect the residential address of the volunteers tested during the winter periods, with the school locations highlighted using open triangles.





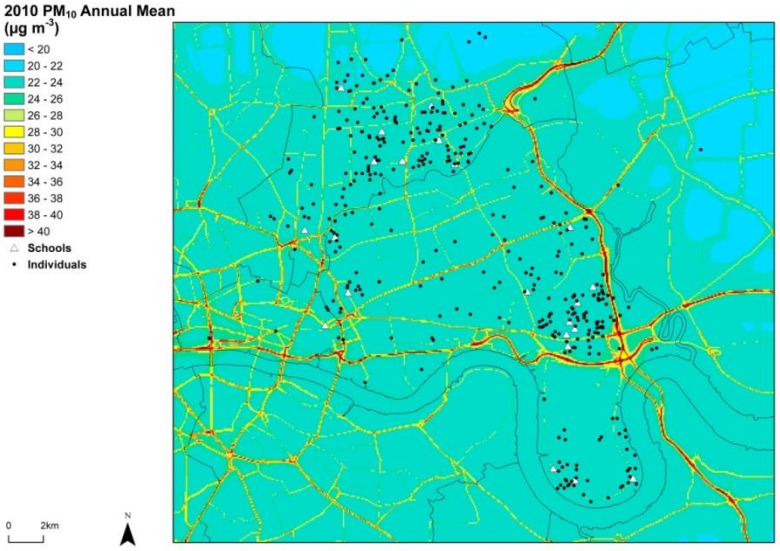
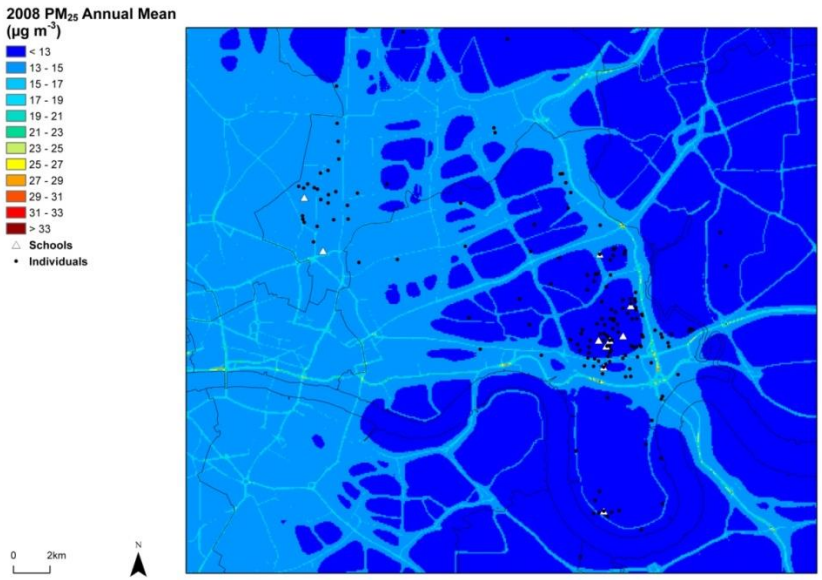
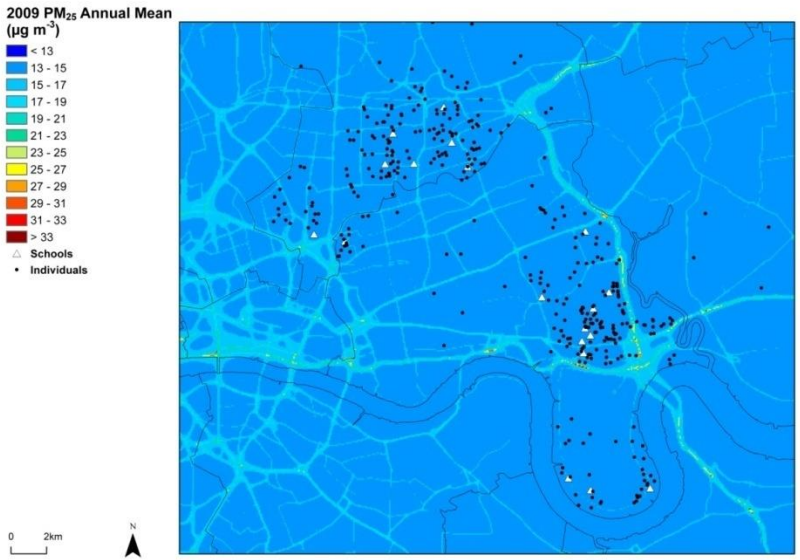


Figure 3.8 Annual PM_{10} maps covering the London boroughs of Hackney and Tower Hamlets for the years 2008 -2010. Individual points reflect the residential address of the volunteers tested during the winter periods, with the school locations highlighted using open triangles.





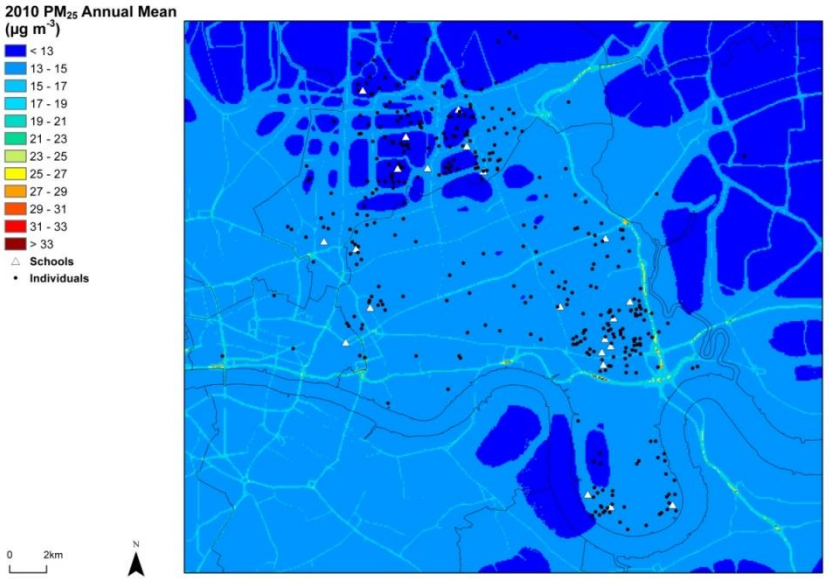


Figure 3.9 Annual $PM_{2.5}$ maps covering the London boroughs of Hackney and Tower Hamlets for the years 2008 -2010. Individual points reflect the residential address of the volunteers tested during the winter periods, with the school locations highlighted using open triangles.

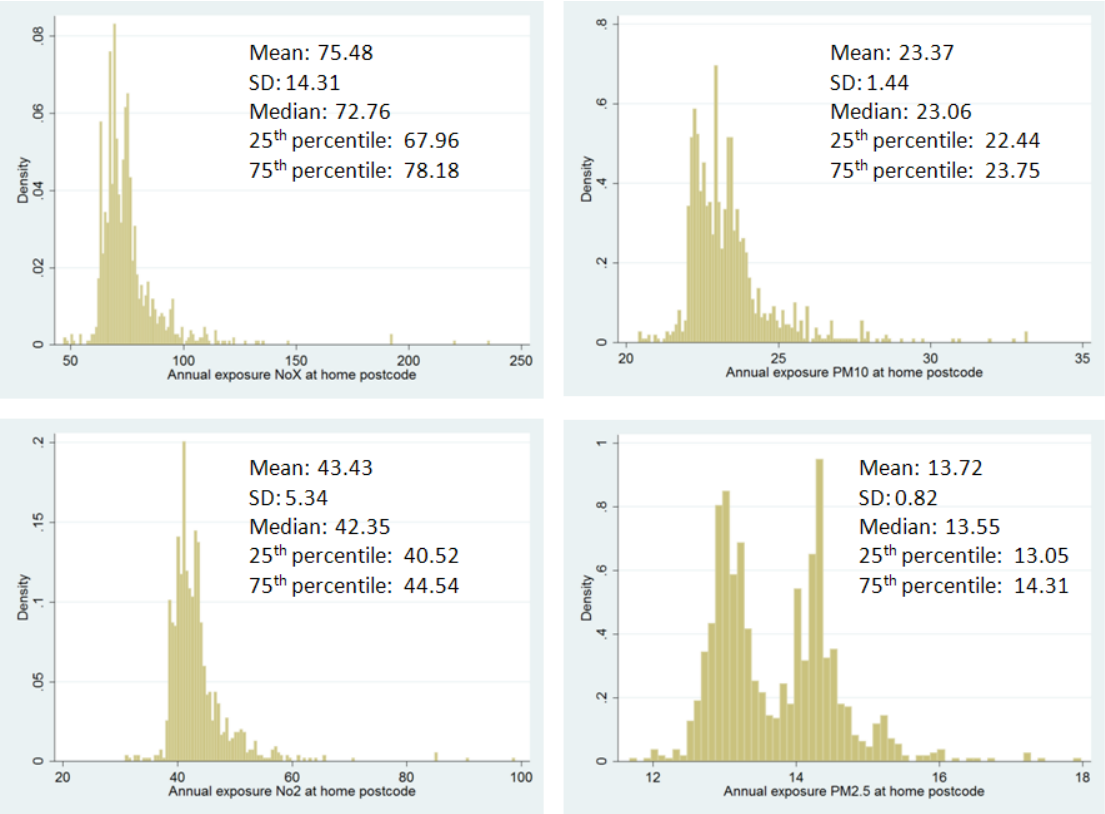


Figure 3.10 Overall distributions of NOx, NO_2, PM_{10} and $PM_{2.5}$ exposure attributions assessed at residential address level.

3.3.3 Symptom Prevalence

The prevalence of asthma symptoms amongst all children, including those with current wheeze are shown in **Table 3.3**. Current wheeze was reported by 11.2% of children and lifetime asthma by 14.4%; however, only 66.7% of those with current wheeze also reported lifetime asthma. Severe wheeze symptoms were reported by 5.5% of children, with 3.9% reporting ≥4 attacks of wheezing in the last year (see **Table 3.3**). However, among children with current wheeze, the prevalence of severe symptoms was high: 29.7% had ≥4 attacks in the last year, 24.3% had ≥1 night/week of disturbed sleep, and 21.6% had wheeze-limited speech, suggesting poor control or lack of asthma diagnosis.

Table 3.3 Prevalence of current symptoms and symptom severity, and of lifetime asthma, hay fever and eczema (%)

	All (n=995)	**Of those with current wheeze** (n=111)	**Of those with current rhinoconjunctivitis** (n=125)	**Of those with current eczema** (n=148)
Asthma				
Current symptoms:				
Wheeze	11.2	-	16.8	11.5
Wheeze with exercise	11.2	63.1	-	-
Dry cough at night	26.2	70.3	-	-
Severe wheeze*	5.5	49.5	-	-
≥4 attacks of wheeze	3.9	29.7	-	-
≥1 night/wk of sleep disturbed by wheeze	3.0	24.3	-	-
Speech limited by wheeze	3.0	21.6	-	-
Asthma ever	14.4	66.7	-	-
Rhinitis				
Current symptoms:				
Runny nose and itchy/watery eyes	12.6	18.9	-	12.8
Runny nose only	24.3	-	-	-
Daily activities disturbed a lot by runny nose	1.9	-	11.2	-
Hay fever ever	27.6	-	66.4	-
Atopic eczema				
Current symptoms:				
Itchy flexural rash	14.9	15.3	15.2	-
≥1 night/wk kept awake by itchy rash	4.6	-	-	18.9
Eczema ever	25.4	-	-	62.2

***** Severe wheeze indicates reported YES to current wheeze, and 1 or more of: ≥4 attacks of wheeze, ≥1 night/wk of sleep disturbed by wheeze, speech limited by wheeze within last 12 months. Prevalence should be read vertically, e.g. 14.4% of all participants reported asthma ever; 66.7% of those who reported current wheeze reported asthma ever.

Current rhinoconjunctivitis was reported by 12.6% of children (see **Table 3.3**); rhinitis (i.e. without symptoms affecting the eyes) was much more common (24.3%). Of those with rhinoconjunctivitis, 11.2% reported that their daily activities were affected a lot by their symptoms, with a similar proportion reporting that they were not at all affected (10.4%). Children with rhinoconjunctivitis most commonly experienced these symptoms during the spring and summer (86.4% during at least one month between March and September), with 6.4% reporting symptoms year round (every month from January to December). Lifetime hay fever was reported by 27.6% of all children, and 66.4% of those with rhinoconjunctivitis.

Current eczema was the most common of the three conditions, reported by 14.9% of children. Of these, 18.9% had ≥1 night/week of disturbed sleep, and for 37.2% their itchy rash had not completely cleared at any time during the last year. Eczema symptoms had most commonly first occurred recently (since the age of 5; 41.2% of those with current eczema), although for 35.8% they had first occurred when the child was <2 years old. Lifetime eczema was reported by 25.4% overall, and 62.2% of those with current eczema.

The overlap between current symptoms of wheeze, rhinoconjunctivitis and eczema is shown in **Figure 3.11**. 28.8% of children (n=287) had symptoms of at least one condition, 7.7% had symptoms of at least two, and 2.0% had symptoms of all three conditions. Among those children with current symptoms, wheeze was most commonly associated with other conditions; 36.9% of children with current wheeze also reported rhinoconjunctivitis, 33.3% also reported eczema, and 18.0% reported symptoms of both the other two conditions.

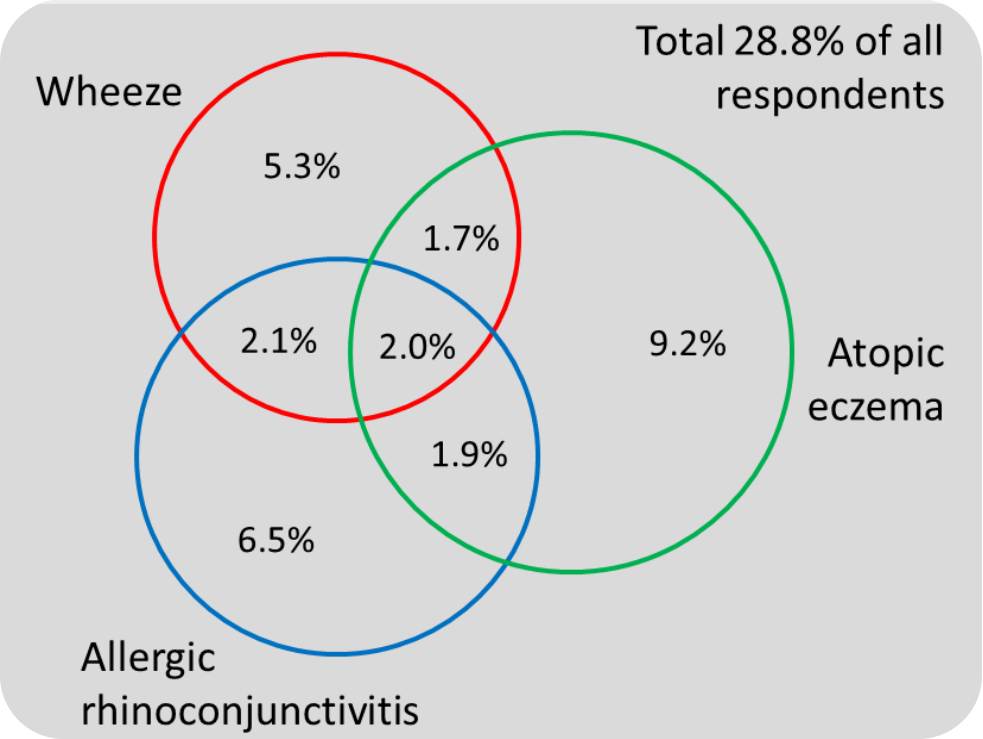


Figure 3.11 Reported prevalence of current respiratory/allergic symptoms. Prevalence (% of total study population, n=995) of wheeze, allergic rhinoconjunctivitis and atopic eczema symptoms within the last 12 months, and co-existence of these symptoms. Overall, 28.8% of children (n=287) had symptoms of at least one condition. Summed prevalences may not match totals shown in Table 4 due to rounding.

3.3.4 Associations with Demographic Variables and Time

The prevalence of current wheeze was negatively associated with age (OR 0.41, 95% CI 0.21-0.82, p=0.011), positively associated with BMI (1.07, 1.00-1.14, p=0.041), was lower in children of Asian background compared with White children (0.46, 0.26-0.81, p=0.008), and was lower in girls than boys (0.47, 0.30-0.74, p=0.001) (see **Table 3.4a**), as

was the prevalence of severe wheeze (0.32,0.12-0.84, p=0.021). Current rhinoconjunctivitis was not associated with any of the demographic variables included in the model; however, the prevalence of current rhinitis was lower in girls compared with boys (OR 0.69, 95% CI 0.50-0.94, p=0.020) and was positively associated with BMI (1.06, 1.01-1.11, p=0.012). The prevalence of eczema was higher in Black than White children (OR 1.85, 95% CI 1.05-3.29, p=0.035). The prevalence of lifetime asthma was negatively associated with age (OR 0.49, 95% CI 0.27-0.90, p=0.021), positively associated with BMI (1.07, 1.01-1.13, p=0.021) and with IMD score (1.02, 1.00-1.04, p=0.042) and was lower in girls than boys (0.62, 0.42-0.93, p=0.019) (see **Table 3.4b**). The prevalence of lifetime hay fever was twice as high in children of all other ethnicities compared with White children (Asian OR 2.08, 95% CI 1.36-3.19, p=0.001; Black 2.09, 1.32-3.30, p=0.002; mixed/other 2.09, 1.22-3.59, p=0.008), while the prevalence of lifetime eczema was lower in children of Asian (OR 0.45, 95% CI 0.28-0.72, p=0.001) and of mixed/other (0.55, 0.31-0.96, p=0.037) ethnic backgrounds compared with White children, and was negatively associated with ETS exposure (0.57, 0.36-0.90, p=0.015).

The prevalence of current respiratory/allergic symptoms and of lifetime hay fever and lifetime eczema did not change over the three-year study period. However, the prevalence of lifetime asthma was lower in Year 3 compared with Year 1 (OR 0.51, CI 0.29-0.90, p=0.020). Overall, 14.4% of children reported lifetime asthma, but year-by-year this percentage decreased significantly from 21.2% in Year 1 to 14.5% in Year 2 to 12.2% in Year 3 (Fishers exact test, p=0.038). However, it should be noted that the study sample size was doubled in Years 2 and 3 compared to Year 1 and included a greater number of schools in Hackney, which may account for some of this variation.

Table 3.4a Odds ratios for associations of personal variables with prevalence of current respiratory/allergic symptoms

	Wheeze (n = 111)	**Rhinoconjunctivitis** (n = 125)	**Eczema** (n = 148)	**Rhinitis** (n = 242)
Age	0.42 (0.21 - 0.82)*	0.83 (0.45 - 1.54)	1.01 (0.55 - 1.84)	1.11 (0.68 - 1.80)
Sex (Female vs. Male)	0.47 (0.30 - 0.74)**	0.88 (0.59 - 1.32)	1.14 (0.77 - 1.68)	0.69 (0.50 - 0.94)*
BMI	1.07 (1.00 - 1.14)*	1.05 (0.99 - 1.11)	1.01 (0.95 - 1.07)	1.06 (1.01 - 1.11)*
ETS exposure	1.14 (0.66 - 1.95)	0.74 (0.43 - 1.29)	0.95 (0.56 - 1.61)	1.01 (0.67 - 1.52)
IMD score	1.01 (0.99 - 1.04)	1.00 (0.98 - 1.02)	0.99 (0.98 - 1.01)	1.00 (0.99 - 1.02)
Ethnicity (Asian vs. White)	0.46 (0.26 - 0.81)**	1.02 (0.60 - 1.75)	1.69 (0.98 - 2.93)	1.19 (0.78 - 1.80)
Ethnicity (Black vs. White)	0.55 (0.30 - 1.01)	0.99 (0.55 - 1.77)	1.85 (1.05 - 3.29)*	0.88 (0.56 - 1.40)
Ethnicity (mixed/other vs. White)	0.73 (0.35 - 1.50)	1.02 (0.50 - 2.09)	0.90 (0.41 - 1.98)	1.22 (0.71 - 2.10)
Study year (Yr 2 vs. Yr 1)	0.97 (0.52 - 1.80)	0.74 (0.42 - 1.32)	0.90 (0.49 - 1.65)	0.97 (0.61 - 1.56)
Study year (Yr 3 vs. Yr 1)	0.60 (0.31 - 1.16)	0.57 (0.32 - 1.04)	1.02 (0.56 - 1.87)	0.69 (0.43 - 1.13)

Data shown as odds ratio (OR) for unit increase in variable unless otherwise stated, with 95% confidence intervals in brackets; OR for each variable adjusted for all other variables shown in table; ETS exposure = positive for urinary cotinine, ajusted for creatinine (CCR > 30ng/mg); * $p<0.05$, ** $p<0.01$.

Table 3.4b Odds ratios for associations of personal variables with prevalence of lifetime (ever having had) asthma, hay fever and eczema

	Asthma (n = 143)	**Hay fever** (n = 275)	**Eczema** (n = 253)
Age	0.49 (0.27 to 0.90)*	0.99 (0.62 to 1.59)	1.41 (0.85 to 2.35)
Sex (Female vs. Male)	0.62 (0.42 to 0.93)*	1.03(0.76 to 1.40)	1.19 (0.86 to 1.64)
BMI	1.07 (1.01 to 1.13)*	1.00 (0.95 to 1.04)	1.03 (0.98 to 1.09)
ETS exposure	1.40 (0.87 to 2.27)	1.11 (0.74 to 1.65)	0.57 (0.36 to 0.90)*
IMD score	1.02 (1.00 to 1.04)*	0.99 (0.98 to 1.01)	0.99 (0.98 to 1.01)
Ethnicity (Asian vs. White)	0.67 (0.39 to 1.16)	2.08 (1.36 to 3.19)**	0.45 (0.28 to 0.72)**
Ethnicity (Black vs. White)	0.92 (0.52 to 1.63)	2.09 (1.32 to 3.30)**	0.76 (0.48 to 1.21)
Ethnicity (mixed /other vs. White)	1.37 (0.72 to 2.57)	2.09 (1.22 to 3.59)**	0.55 (0.31 to 0.96)*
Study year (Yr 2 vs. Yr 1)	0.64 (0.37 to 1.10)	1.11 (0.68 to 1.81)	1.43 (0.83 to 2.48)
Study year (Yr 3 vs. Yr 1)	0.51 (0.29 to 0.90)*	1.25 (0.76 to 2.04)	1.22 (0.70 to 2.13)

Data shown as odds ratio (OR) for unit increase in variable unless otherwise stated, with 95% confidence intervals in brackets; OR for each variable adjusted for all other variables shown in table; ETS exposure = urinary cotinine concentration adjusted for creatinine concentration, * $p<0.05$, ** $p<0.01$.

3.3.5 Associations with Air Pollution

Air pollutant concentrations across the study area changed little during the first three years of the study (**Table 3.5**). Concentrations were assessed between 2008-2010 based on data obtained from sites within the London Air Quality Network (Environmental Research Group) in and surrounding the study area. For $PM_{2.5}$ all sites across London were included, due to the more limited monitoring of the pollutant with London and the UK. For both NOx and NO_2 there was no evidence of a reduction at either urban background or roadside locations. Similarly, $PM_{2.5}$ and PM_{10} concentrations were unaltered at background sites over the first three years of the LEZ. Some evidence of a roadside decrease in PM_{10} was noted between 2008 and 2009, but only when measurements based on Tapered Element Oscillating Microbalance (TEOM) masses were employed ($p=0.027$), with a similar trend in $PM_{2.5}$ TEOM concentrations ($p=0.06$). Similar reductions were not apparent

between 2009 and 2010. These small changes were not reflected in the modelled exposure attributions based on the children's residential addresses (**Table 3.6**).

Table 3.5 Measured annual mean pollutant concentrations ($\mu g/m^3$) at selected background and roadside sites surrounding the study area

	Site type	NOx	NO_2	PM_{10}	$PM_{2.5}$ (FDMS)	$PM_{2.5}$ (TEOM)
Year 1 (2008)	Background	68.4 ± 19.8 (40.2-101.5, n=8)	41.1 ± 8.7 (26.0-52.6, n=8)	22.0 ± 1.5 (20.7-24.9, n=8)	No data	10.5 ± 0.4 (10.3-11.0, n=3)
Year 2 (2009)	Background	66.1 ± 22.1 (36.9-105.2, n=8)	40.8 ± 10.1 (23.5-57.2, n=8)	22.5 ± 1.91 (20.4-25.8, n=6)	14.7 ± 1.8 (13.2-17.6, n=6)	10.3 ± 0.9 (9.7-11.6, n=4)
Year 3 (2010)	Background	67.4 ± 23.4 (36.8-100.6, n=6)	40.5 ± 10.5 (24.3-48.6, n=6)	21.9 ± 0.5 (21.5-22.7, n=5)	14.7 ± 1.4 (12.8-16.5, n=7)	10.4 ± 1.4 (9.4-12.5, n=4)
Year 1 (2008)	Roadside	135.9 ± 41.1 (89.8-223.0, n=13)	57.2 ± 10.8 (43.7-70.3, n=13)	28.7 ± 6.1 (20.3-40.9, n=11)	17.1 ± 1.2 (16.0-18.5, n=4)	14.9 ± 3.3 (11.5-20.2, n=8)
Year 2 (2009)	Roadside	131.9 ± 47.4 (87.2-255.7, n=14)	56.8 ± 12.1 (44.2-82.3, n=14)	26.9 ± 5.1* (20.5-36.9, n=11)	15.1 ± 3.1 (11.7-19.0, n=4)	14.2 ± 3.0 (11.1-18.6, n=7)
Year 3 (2010)	Roadside	131.7 ± 40.2 (82.7-220.3, n=14)	56.7 ± 11.7 (41.8-74.0, n=14)	26.2 ± 3.6 (19.9-32.6, n=12)	17.6 ± 1.7 (15.1-19.9, n=5)	13.9 ± 2.1 (11.4-16.4, n=4)

Annual mean ± SD pollutant concentrations (range, number of sites included), based on sites within and surrounding the London boroughs of Hackney and Tower Hamlets. Annual means were calculated for each of the included sites with greater than 75% data capture across the given year. PM_{10} is expressed as reference equivalent concentration, based on the correction of Tapered Element Oscillating Microbalance (TEOM) masses for the loss of volatiles components using the volatile correction method (Green et al. 2009) and combined with measurements made using the Filter Dynamics Measurement System (FDMS). As no empirical method is available for correcting $PM_{2.5}$ TEOM masses the TEOM and FDMS mass concentrations are given separately. The average annual mean concentrations across each of the site types was compared using a one way ANOVA, with post hoc testing performed using paired t-tests: *, significant reduction ($p<0.05$) in pollutant concentration between 2008 and 2009.

Table 3.6 Modelled annual pollutant attributions for the children recruited into the study, based on their residential address, by year and averaged across study period, μg/m^3

	N	NOx	NO_2	PM_{10}	$PM_{2.5}$
Year 1 (2008)	131*	71.0 ± 10.6 (61.5-122.8)	41.9 ± 3.9 (38.4-61.0)	22.9 ± 1.3 (21.7-29.8)	13.0 ± 0.5 (12.5-16.0)
Year 2 (2009)	418*	79.3 ± 13.1 (50.8-235.4)	44.9 ± 4.9 (32.9-98.9)	23.9 ± 1.2 (20.7-32.7)	14.4 ± 0.5 (12.6-18.0)
Year 3 (2010)	444	73.8 ± 16.4 (47.1-220.8)	42.7 ± 6.1 (30.9-90.8)	23.0 ± 1.5 (20.4-33.2)	13.3 ± 0.6 (12.0-17.3)
Years 1-3	993	75.7 ± 14.8 (47.1-235.4)	43.5 ± 5.5 (30.9-98.9)	23.4 ± 1.5 (20.4-33.2)	13.7 ± 0.8 (12.0-18.0)

Values are mean ± SD (range); * air pollution data were not available for two participants (one in each of Years 1 and 2 of the study) because their residential addresses were outside the Greater London area.

The prevalence of current rhinitis was positively associated with annual mean NOx (OR 1.01, 95% CI 1.00-1.02, p=0.033), NO_2 (1.03, 1.00-1.06, p=0.034), PM_{10} (1.16, 1.04-1.28, p=0.006) and $PM_{2.5}$ (1.38, 1.08-1.78, p=0.011) (**Figure 3.12**, ORs for an increase of 1 μg/m^3). The prevalence of current wheeze, rhinoconjunctivitis and eczema symptoms and of lifetime asthma, hay fever and eczema ever was not associated with air pollution levels.

3.3.6 Associations with Asthma 'Genotype' and 'Phenotype'

The prevalence of lifetime asthma was higher with the homozygous genotype 'TT' SNP in GSDMA, compared with the wild type genotype 'CC' (OR 1.99, 95% CI 1.15-3.45, p=0.013), whereas no association was found for the heterozygous genotype CT (vs. genotype CC, see **Table 3.7**). In contrast, the prevalence of lifetime asthma was lower with the homozygous genotype 'TT' SNP in GSDMB, compared with the wild type genotype

'CC', suggesting a protective effect (OR 0.44, 95% CI 0.19 to 0.99, p=0.046), while no association was found for the heterozygous genotype CT. The prevalence of current wheeze was not associated with any of the SNPs in GSDMA or GSDMB.

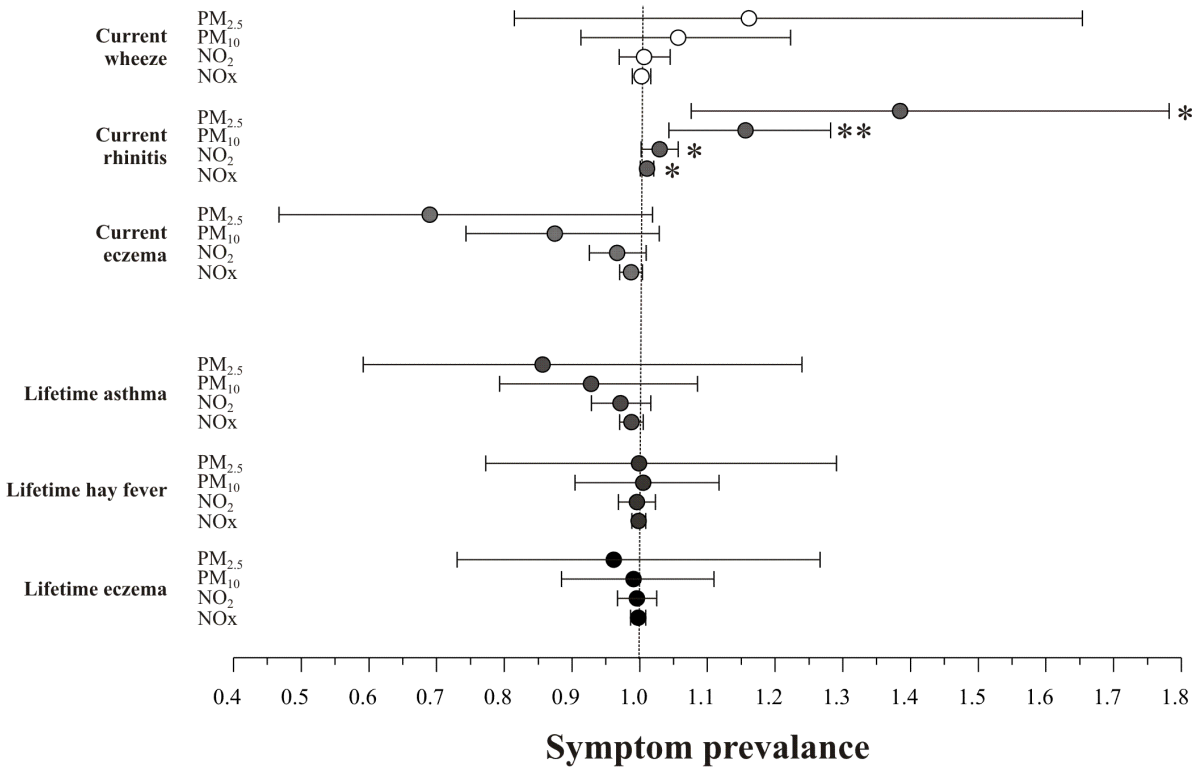


Figure 3.12 Exposure to air pollution as a risk factor for current and lifetime respiratory/allergic symptoms. Risk expressed as adjusted odds ratio (OR 95% CI), calculated by multiple logistic regression analyses adjusted for year of study, age, sex, BMI, socio-economic deprivation (IMD score), and exposure to environmental tobacco smoke exposure, with a random effect for school. Single-pollutant models were calculated for each air pollutant. Odds ratios are for unit increase in pollutant, in $\mu g/m^3$. Current symptoms defined as within the last 12 months; lifetime conditions defined as 'having ever had' asthma, hay fever or eczema (see methods for full definitions). Vertical dotted line indicates null (OR = 1). Horizontal lines indicate 95% confidence intervals of odds ratios. * p<0.05, **p<0.01.

Table 3.7 Odds ratios (95% confidence intervals) for associations of gasdermin A and B SNPs with prevalence of current wheeze and lifetime asthma (among all respondents).

	Wheeze (n=111)	**Asthma** (n=143)
GSDMA rs3894194 - Arg18Gin (C/T)		
CC	Reference (n=38)	Reference (n=40)
CT	0.70 (0.42 to 1.15) (n=40)	1.09 (0.69 to 1.73) (n=58)
TT	1.28 (0.71 to 2.32) (n=23)	1.99 (1.15 to 3.45)* (n=32)
GSDMB rs2305480 - Pro298Ser (C/T)		
CC	Reference (n=50)	Reference (n=72)
CT	0.93 (0.57 to 1.50) (n=43)	0.81 (0.52 to 1.24) (n=50)
TT	0.55 (0.23 to 1.30) (n=8)	0.44 (0.19 to 0.99)* (n=9)

* $p<0.05$, n number shown in brackets below; of those with wheeze, 10 subjects were not genotyped for GSDMA or GSDMB; of those with asthma, 13 subjects were not genotyped for GSDMA and 12 subjects were not genotyped for GSDMB.

A subset of children with poorly-controlled or undiagnosed asthma was identified as those with a significant BDR. The prevalence of current wheeze and of lifetime asthma was positively associated with BDR when expressed as a continuous variable (OR 1.07, 95% CI 1.03-1.12, p=0.001 and 1.05, 1.02-1.09, p=0.014, respectively), and when expressed as a yes/no variable (current wheeze 2.51, 1.44-4.39, p=0.001; lifetime asthma 1.91, 1.12-3.25, p=0.018). As such, this objective measurement of airway hypersensitivity helps to validate the questions on wheeze and asthma as accurately identifying children with respiratory symptoms.

Including the annual PM_{10} level in the models revealed no effect modification of either gasdermin SNPs or airway hypersensitivity; PM_{10} level had no effect on the prevalence of current wheeze or lifetime asthma regardless of asthma 'genotype' or 'phenotype'.

3.4 DISCUSSION

The present study addressed the allergic and respiratory symptoms of children residing with London's LEZ, over the first three years of its operation. Of the symptoms examined, only current rhinitis was positively associated with annual modelled NOx, NO_2, PM_{10} and $PM_{2.5}$ exposures and we observed no evidence of a reduction in symptom prevalence over the study period. However, we also failed to observe the predicted improvements in air quality. A high prevalence of severe asthma symptoms was also observed among children with current wheeze, with low numbers reporting lifetime asthma, indicating poor asthma control, under-diagnosis or under-reporting.

3.4.1 Comparison with Previous Studies

Only one other study has examined the potential health benefits of introducing a Low Emission Zone in an urban area (Cesaroni et al. 2012). Cesaroni and co-workers investigated the effects of two LEZs in Rome, Italy; however, they did not measure any health parameters at the individual level; rather they estimated 'years of life gained' using concentration-response functions derived from cohort studies in the USA and Europe (Cesaroni et al. 2012). Using this approach, they reported that residents living along busy roads gained 3.4 days/person as a result of the zones being introduced, based on the modelled projections of the impact of the scheme. Notably, no formal confirmation of the impact of these LEZs upon air quality in Rome was performed in this study and indeed

subsequent examination of the air monitoring data within the zone failed to demonstrate the modeled reductions in fine particulate.

The observations over the first three years of the LEZ are consistent with those of Anderson and colleagues (Anderson et al. 2010), who used data from ISAAC Phase I and modelled city-level residential PM_{10} data to explore associations with the prevalence of asthma, rhinoconjunctivitis and eczema. This analysis included data for over 500,000 children from 51 countries. Weak negative associations were found for city-level PM_{10} and various outcomes, while for country-level PM_{10} (for 24 countries with >1 centre) most associations were weakly positive. The authors concluded that urban background PM_{10} had little or no association with the prevalence of childhood asthma, rhinoconjunctivitis or eczema, either within or between countries (Anderson et al. 2010).

Later findings from ISAAC Phase III, have however, shown a positive relationship between symptoms of asthma and allergic disease and traffic-related pollution (Brunekreef et al. 2009). An environmental questionnaire, including a question on frequency of truck traffic on the street of residence, was included in this phase. A highly significant exposure-response relationship between truck traffic frequency and respiratory/allergic symptoms was reported, with the strongest relationship for symptoms of severe asthma (Brunekreef et al. 2009).

Other studies using self- or parent-reported symptom data from questionnaires have shown positive associations between respiratory/allergic symptoms and air pollution. Frequency of (self-reported) heavy truck traffic on the street of residence was positively correlated with the prevalence of wheezing and allergic rhinitis among schoolchildren in Bochum, Germany (Weiland et al. 1994). Zuraimi et al. (2011) found significant dose-response relationships between (self-reported) traffic density and asthma and rhinitis symptoms in pre-school children in Singapore, which were stronger among children sleeping in non-air-conditioned homes. A number of studies have used objective measures of traffic density as a proxy for air pollution exposure. In Munich, Germany, an increase of

25,000 cars daily passing through the school district on the main road was associated with a significant increase in the prevalence of recurrent wheezing with dyspnoea and of recurrent dyspnoea in 9-11 year-old schoolchildren (Wjst et al. 1993). Children living on busy streets (based on modelled NO_2 concentrations) reported significantly higher prevalence of wheeze than children living on quiet streets in Haarlem, The Netherlands (Oosterlee et al. 1996). Dales et al. (2009) found that roadway density was significantly associated with wheeze (OR 1.23), wheeze with dyspnoea (OR 1.27) and asthma (OR 1.08) in elementary schoolchildren in Windsor, Ontario. In this study, the measure of traffic exposure was only resolved to the level of the child's neighbourhood, as road density was calculated from the summed length of all roadways within a 200m radius of the home postcode.

Other studies have used data from ambient air pollution monitoring sites to assign exposure at the individual level. Using a validated three-year dispersion model, exposures were assigned to school addresses for 9-11 year-old schoolchildren in the French Six Cities Study (Penard-Morand et al. 2010). PM_{10} levels were associated with asthma, eczema and allergic rhinitis, and NOx was associated with asthma. A number of studies from The Netherlands have shown significant associations between proxy or actual measures of traffic pollution and respiratory symptoms, but only among subgroups of children with chronic respiratory symptoms (van der Zee et al. 1999), or bronchial hyper-responsiveness and allergic sensitization (Boezen et al. 1999; Janssen et al. 2003).

Differences in study design may explain some of the inconsistencies in reported associations between air pollution and respiratory/allergic symptoms, e.g. differences in study area, measured pollutants or proxy measure of pollutant exposure, wording of symptom questions or study population (Hoek et al. 2012). A recently published meta-analysis has attempted to address these issues by combining the original data from 11 studies on PM_{10} and respiratory symptoms, including data on >45,000 children from 12 countries (Hoek et al. 2012). This study found that symptoms of wheeze and asthma diagnosis were not associated with PM_{10}, although there was a weak association with hay

fever diagnosis. An earlier meta-analysis by the same group examined associations between NO_2 and respiratory symptoms and included original data from 5 studies (24,000 children from 5 countries). Pattenden et al. (2006) found no overall evidence of associations between ambient NO_2 and asthma diagnosis or wheeze.

In the current study, NO_2 and PM_{10} concentrations were within a similar range to most of the studies cited above, at least for urban areas (Boezen et al. 1999; Hoek et al. 2012; Pattenden et al. 2006; Penard-Morand et al. 2010; van der Zee et al. 1999). Interestingly, for 84% of the children in our study, the annual mean residential NO_2 level was above the London air quality objective (and World Health Organization, WHO, guideline) of 40 $\mu g/m^3$. The objective for PM_{10}, also 40 $\mu g/m^3$, was not exceeded for any of the children. However, air quality guidelines published by WHO in 2005 (WHO 2006) recommended an annual mean of 20 $\mu g/m^3$ for PM_{10}, which was exceeded for all the children in the current study.

Compared with symptom prevalence data for Great Britain from ISAAC Phase I, prevalence among our study population was considerably lower; a third of schoolchildren reported wheeze in the past year (vs. 11.2% in our study) and a fifth recalled a diagnosis of asthma (vs. 14.4% in our study) (Kaur et al. 1998). However, data from ISAAC UK was for 12-14 year-olds, only a small number of the schools were located in London and most were in non-metropolitan areas. In comparison, our data are strikingly similar to those for the 8-10 year-old group from a survey of asthma prevalence and severity in Great Britain in 1992 (Strachan et al. 1994). The prevalence of current wheeze was somewhat higher than in our study (15% overall in 5-17 year-olds), but a larger proportion of children with current wheeze reported frequent asthma attacks in our study – almost one in three, vs. one in six (Strachan et al. 1994). These authors (Strachan et al. 1994) found that wheezing and symptom severity was associated with lower socioeconomic status, which we did not; however, this could be due to a lack of range within the IMD 2010 score among our study population, since the level of deprivation is high throughout Tower Hamlets and Hackney

(Department for Communities and Local Government 2011). Strachan et al. (1994) stated that wheezing illness 'is clearly a major public health problem'; given the similarity to the current data this still appears to be the case 20 years later, at least in east London.

Although over 10% of the children in our study reported current wheeze, of those, only two-thirds reported lifetime asthma. Similarly, Strachan et al. (1994) reported that asthma had been diagnosed in only 54% of the wheezy children in that study. The apparent disparity between reports of current wheeze symptoms and having asthma could reflect under-diagnosis or GPs not specifically using the term 'asthma', or possibly that medical advice/care has not been sought regarding these symptoms. We also found a high prevalence of severe asthma symptoms among those with current wheeze, which could again indicate a lack of diagnosis or medical care, or poor symptom control. We included bronchodilator response in our analysis, and found an elevated response to be associated with current wheeze and lifetime asthma, which would appear to confirm this finding using an objective measure of airway hypersensitivity. A published 'factsheet' regarding asthma in the borough of Tower Hamlets acknowledges that under-diagnosis or recording issues may explain why the prevalence is lower than in London and England (Tower Hamlets Council 2011), but also indicates that the hospital admission rate for asthma is higher than the London average, when adjusted for age, which suggests that poor asthma control may be an issue – especially for children – in this area of London.

A recent study investigated the association of GSDMA and GSDMB variants with asthma in Korean children and found that polymorphisms in these genes may be associated with asthma susceptibility and intermediate phenotypes of asthma, such as bronchial hyper-responsiveness in children with asthma (Yu et al. 2011). These results suggested an important role for gasdermin genes in the development of childhood asthma. We report here a positive association between the same SNP in GSDMA and lifetime asthma, which concurs with its apparent involvement in childhood-onset asthma. The SNP in GSDMB analysed in our study was different from that in Yu et al. (Yu et al. 2011) and has been

found to have a negative association with childhood-onset asthma (Moffatt et al. 2010). Similarly, our results indicate a negative association with lifetime asthma, suggesting a possible protective effective of this polymorphism in GSDMB.

3.4.2 Strengths and Limitations

This study was designed to target areas with high levels of air pollution, where the greatest air quality improvements were predicted to result from the LEZ. It was a population survey based on school classes, including children from a wide diversity of ethnic groups, which have often been excluded in previous studies. We employed a validated, widely-used questionnaire to measure respiratory/allergic symptoms, and linked these parent-reported responses to a large number of biological measurements, including lung function, and genotype of the gasdermin genes involved in childhood onset asthma. In addition we employed a validated, high-resolution air pollution model, providing exposure measurements at the level of residential address, over a period that encompasses the recent dieselisation of the London vehicle fleet and the resulting changes in the air pollutant mix.

A limitation of the type of parent-reported symptom data used herein is the possibility of recall bias or exaggeration of symptoms, especially given that parents were told we were investigating the effects of the LEZ on children's health. However, given that the symptom prevalences reported here were lower than those from the ISAAC UK study (Austin et al. 1999; Kaur et al. 1998) exaggeration does not appear to have been a problem. Some inconsistencies in the data may be explained by language issues since English was probably not the first language for a large number of respondents. The term 'wheeze' does not translate exactly into all languages, as noted by the ISAAC study (The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee 1998). We administered the ISAAC questionnaire in English, as it was beyond our resources to get

translations made and validated in all the languages required, or to verbally administer the questionnaire to speakers of Sylheti, which has no written form.

There are several reasons why the predicted air quality improvements from the LEZ have not occurred, including the delay in implementing phase III (originally scheduled for October 2010, which would have applied the Euro III PM standard to light goods vehicles), the increasing proportion of diesel cars within the fleet (Cames and Helmers 2013), and evidence that NOx emissions from newer diesel engines (Euro 3-5) have not fallen as predicted by the current emission inventories (Carslaw et al. 2011). As the delayed phase III was introduced with phase IV (requiring heavy goods vehicles, buses and coaches to meet the Euro IV PM standard) in January 2012, it is feasible that the predicted improvements may only become apparent in the subsequent years, although this is critically dependent on the emission technologies delivering under real-world driving conditions. Our study is on-going for a further three years, permitting continued analysis of any LEZ-related effects on air quality and subsequent changes in the prevalence of respiratory/allergic symptoms.

3.4.3 Conclusions

Over the first three years of the LEZ there was no evidence of a decrease in the respiratory/allergic symptoms of schoolchildren living in east London. Consistent with previous studies current rhinitis was positively associated with traffic-related air pollutants, while symptoms of other respiratory/allergic conditions were not. Associations were found between lifetime asthma and SNPs in gasdermin genes, which have previously been shown to be associated with child-onset asthma. Positive associations were also noted between current wheeze and lifetime asthma and significant bronchodilator responses, an objective measure of airway hypersensitivity. An additional unexpected finding was the high prevalence of severe asthma symptoms among children with current wheeze, and high

proportion of these who did not report having asthma. These findings indicate possible issues with under-diagnosis of asthma, under-use of health care services, and/or lack of control of asthma symptoms. Exploring these issues in more depth should help GPs and other healthcare workers to understand how best to ensure these children get the most appropriate treatment for their symptoms.

Chapter 4

Impact of Traffic-Related Pollution on Respiratory Function in Children Living within London's Low Emission Zone

4.1 INTRODUCTION

As outlined previously in the present study I hypothesised that year-on-year improvements in air quality following the introduction of the LEZ would results in corresponding improvements in children's lung function and airway respiratory symptoms. In Chapter 3 I demonstrated that whilst there was a relationship between annual pollutant exposures and rhinitis, there was no year-on-year change in symptoms over the first three years of the schemes operation. This was argued to reflect the absence of quantifiable improvements in air quality, related to the failure of Euro IV and V engines to produce their envisaged emission improvements in the real world. In this study I present an analysis of the relationship between lung function and air pollutant exposures in children (8-9 years of age) living within the LEZ for the first three years of its operation (post the implementation of phases 1 and 2, but prior to phase 3, which was introduced in January 2012). We focused on children because their developing respiratory systems appear particularly sensitive to adverse air quality (Gauderman et al. 2007; Schultz et al. 2012). Children are especially vulnerable to traffic emissions for a number of reasons: First, lung development is still occurring in early life and therefore injury in this crucial period may have long term consequences for respiratory health; in particular, failure to attain maximum potential lung function means children have a greater likelihood of early chronic lung disease in adulthood. Second, children are likely to spend a greater period of time outdoors than adults, exposing them to higher doses of pollutants.

Numerous studies have already demonstrated evidence of impaired lung function (cross sectionally) and growth (longitudinally) in young populations exposed to urban and/or traffic pollution (Jedrychowski et al. 1999; Peters et al. 1999; Horak et al. 2002; Gauderman et al. 2004, 2007; Rojas-Martinez et al. 2007; Oftedal et al. 2008; He et al. 2010; Roy et al. 2012; Schultz et al. 2012). Of these studies the most coherent evidence was obtained from the Californian Children's Health Study, which reported associations between community-average concentrations of acid vapour, $PM_{2.5}$ and elemental carbon and diminished lung function development in children aged 10–18 years, (Gauderman et al. 2000; Gauderman et al. 2002; Gauderman et al. 2004). Of particular relevance to this study, a subsequent follow up of children who moved between the 12 communities from high to less polluted areas appeared to show they experienced diminished deficits in lung growth (Avol et al. 2001). Similar associations have been reported in Poland and Austria (Jedrychowski 1999; Horak 2002). Whilst these studies focused on pollutant concentrations, similar associations have been noted in relation to simple proxies of traffic exposures, with evidence of impaired lung growth in children living within 500m of a freeway, compared with children living at least 1,500 m from a major road (Gauderman et al. 2007). Whilst not all studies addressing the adverse effects of air pollutants on lung function have demonstrated similar reduction in lung volume, or flow variables (Dockery et al. 1989; Hirsch et al. 1999; Nicolai et al. 2003; Hoek et al. 2012), the overall consensus view supports a negative relationship (Gotschi et al. 2008; HEI 2010). These early life decrements in lung volume (in particular FEV_1) though small are highly important as they have been shown to be a strong determinant of life expectancy (Higgins and Keller 1970; Ashley et al. 1975; Hole et al. 1996), as well as for respiratory effects in later life (Turnovska et al. 2009; Moshammer et al. 2006; Gabriele et al. 2008; Clark et al. 2010).

Traffic emissions are the major source of airborne pollutants in urban areas, with the combustion aerosol containing numerous harmful substances such as benzene, particulate matter (PM) and oxides of nitrogen; many of which have been shown to induce oxidative stress in vivo (Valavanidis et al. 2009). In particular, diesel vehicles contribute significantly

to road side fine particle concentrations, elemental carbon, particle number concentrations and oxides of nitrogen, especially in Europe where they represent a major source of polycyclic aromatic hydrocarbons (PAHs) and their derivatives like quinones and carboxylic and nitro- PAHs in the airshed (Shinyashiki et al. 2009). Respirable PAHs are genotoxic, causing DNA adduct formation resulting in oxidative DNA damage and are associated with an increased risk of cancer (Shimada et al. 1992). The metabolism of the PAHs is initiated by its binding to the aryl hydrocarbon receptors, followed by the transformation by phase I enzymes, cytochrome P450 (CYP)1A1 to highly reactive diol epoxide, which is subsequently detoxified by phase II enzymes, including epoxide hydrolases and glutathione S-transferases (e.g. GSTM1 and GSTP1). The capacity of an individual to up-regulate these detoxification pathways in response to pollutant challenge has therefore been proposed as a determinant of an individual's susceptibility to these inhaled xenobiotics (Sandstrom and Kelly 2009).

To date the majority of studies examining the influence of genetic variants in these genes on air pollutant, respiratory health associations have focused on the phase II xenobiotic enzymes, often incongruously referred to as antioxidants in the literature. These studies (reviewed in **Chapter 1**, **section 1.3.2.1**, and summarized in **Table 1.3** and **Table 1.4**) have predominately focused on polymorphisms in members of the GST supergene family, with GSTM1 (Buthbumrung et al. 2008; David et al. 2003; Hong et al. 2007; Islam et al. 2009; Lee et al. 2008; Li et al. 2006; Piacentini et al. 2010; Reddy et al. 2012; Romieu et al. 2004, 2006; Salam et al. 2007; Tung et al. 2011), GSTP1 (Avogbe et al. 2005; Alexeeff et al. 2008; Chen et al. 2007; Romieu et al. 2006) and GSTT1 (Avogbe et al. 2005; Piacentini et al. 2010) being the most widely studied; followed by NQO1 (Nebert et al. 2002; Bergamaschi et al. 2001; Corradi et al. 2002; Chen et al. 2007; David et al. 2003; Avogbe et al. 2005), microsomal epoxide hydrolase (Salam et al. 2007; Tung et al. 2011; Manini et al. 2006; Binkova et al. 2007; Novotna et al. 2007), and more recently enzymes involved in glutathione synthesis: glutathione synthase ligase, GSSG reductase, and glutamate cysteine (Breton et al. 2011). Only relatively few of these studies have examine

polymorphisms in these genes on lung function, air pollutant interactions, with the majority focused on symptom prevalence and markers of exhaled or systemic oxidative stress (**Table 1.3** and **Table 1.4**). Those studies examining lung function have focused on genetic variants in GSTM1 (null), GSTP1(Ile105Val) and NQO1 (Pro187Ser) in adults and children with high ambient ozone exposures, demonstrating either enhance decrements in FEV_1 (Bergamaschi et al. 2001; Alexreef et al. 2008; Romieu et al. 2004), decreased FEF_{25-75} (Chen et al. 2007), or increased difficulty in breathing (Romieu et al. 2006). In studies examining $PM_{2.5}$ or PM_{10}, the impact of these genetic variants on lung function are less clear (Hong et al. 2007). There is consequently an absence of studies investigating the influence of phase II xenobiotic genes on lung function changes related to primary roadside emissions. Many of the studies cited above are also based on very small group sizes, with the effect often only observed as an interaction term between SNPs, which the studies are seldom adequately powered to address (as reviewed in Minelli et al. 2010). There are also significant limitations on the exposure attributions used in these studies, with the assessments often limited to data from a single or restricted subset of monitoring sites, and with limited or no attempt made to demarcate acute from chronic effects.

In the present study I therefore examined how SNPs in a panel of selected genes: GSTM1, GSTP1 and NQO1 (selected based on the literature review present in **Table 1.3** and **Table 1.4**), and referred to throughout as the confirmatory panel, influenced lung function air pollutant interactions. For this assessment I employed post bronchodilator lung function and exposure attributions based on the residential address expressed over various temporal domains: acute (24 hour pre-health assessment), sub-chronic (7 day average) and chronic (annual) exposures, based on a high resolution dispersion model and a NOWCAST scaling factor, developed for this project to provide spatially refined acute exposure estimates. In addition to the confirmatory panel of SNPs previously reported to influence pollutant, lung function interactions, I also examined an exploratory panel based on polymorphisms in the aryl hydrocarbon receptor (AhR) and CYP1A1. Most of the previous studies examining phase I xenobiotic genes have focused on the occupational exposures of

workers to pollutants such as PAHs (Chen et al. 2007; Bin et al. 2008) and cigarette smoke (Figueroa et al. 2008; Chen et al. 2009). To the best of my knowledge, there is no literature on the effect of genetic variations of the AhR and CYP1A1 on ambient air pollution, lung function interactions.

My primary objective in this study was to quantify the impact of urban air pollution on indices of lung function and to examine whether any such changes were the result of long or short-term exposures to NOx, NO_2, PM_{10} or $PM_{2.5}$. As a secondary objective, I also examined whether variation in genes related to the phase I and II xenobiotic metabolism of PAHs impacted upon lung function and the interaction between pollutant exposures and indices of lung capacity (FVC and FEV_1). I examined two panels of SNPs, the first confirmatory, based on previously reported gene interactions between lung function and air pollutant exposures, focused on polymorphisms in glutathione-S-transferases and NADPH:quinone reductase. The second panel was exploratory, examining SNPs in CYP1A1 and AhR due to their role in PAH metabolism and previous in vitro and in vivo evidence that these proteins are involved in modulating the toxicity of diesel exhaust particles. I also examined whether there were any year-on-year improvements in lung function related to residence within the LEZ, based on the assumption that accrued improvements in air quality would be associated with better respiratory health. As stated in the previous chapter however, given the absence in of a change in the pollutants examined over the first three years of London's LEZ, this study should be viewed predominately as establishing a baseline understanding of the underlying pollutant lung function interactions prior to the implementation of phases III and IV of the LEZ in January 2012.

4.2 METHODS

4.2.1 Study Design and Study Population

Details of the study design and general health assessment were as provided in **Chapter 3** (**sections 3.2.1** and **3.2.3**).

4.2.2 Lung Function Assessments

Children's respiratory function was assessed by spirometry (Microlab, Micromedical, Carefusion), performed by trained investigators according to ATS-ERS guidelines (2005) with baseline and post-bronchodilator measurements, following salbutamol 400 µg administered by large volume spacer. Before each measurement, volume calibration with a 3L syringe was undertaken. A sterile disposable filter/ mouthpiece was attached to the spirometer for each child and the equipment wiped with alcohol wipes between subjects. A maximum of 10 attempts was normally made (exceptionally, more if needed) until three acceptable and two repeatable attempts were attained. Each spirometry measurement aimed to obtain three acceptable and two repeatable attempts both pre and post bronchodilator. A short acting bronchodilator (salbutamol) was administered after baseline spirometry. Four 100 microgram actuations were given from a metered dose inhaler (MDI) through a spacer device (Volumetric). The children took 4 tidal breaths through the volumetric spacer after each actuation. A minimum of 15 minutes later post bronchodilator spirometry took place.

Quality control was based on the ATS/ERS guidelines (2005), modified for children (Praud 2004). Additional acceptability requirements were: rapid onset of expiration, high well defined peak flow and a clear plateau on volume-time curve together with no evidence of cough, glottis closure or leak during the manoeuvre (from Asthma UK/ Growing Lungs

Guidelines). For quality control and reporting purposes spirometry results were uploaded from the 3 study spirometers to a Carefusion program "Spirometry" where individual inspection of efforts was undertaken in detail. Reporting of results was according to ATS-ERS Guidelines and the best overall individual effort from each child both pre and post salbutamol was selected as the highest value of FEV_1 and FVC, together with FEF_{25-75} from the best combined effort, reported from technically acceptable data, which met reproducibility requirements of 0.15 L agreement for FEV_1 and FVC for 2 efforts. Data were then extracted into Excel and individual efforts were cross checked with original data before being exported into the main study Access database. Raw data were transformed into *Z* (or standard deviation) scores using most the "All-Age" equations by Stanojevic et al. (2009) available for white subjects aged 3 to 80 years of age.

4.2.3 Urine Collection and Analysis

Spot urine samples were collected and placed on ice for the duration of the visit, then frozen and stored at -80°C on return to the laboratory. Urine cotinine concentrations were determined using a commercial microplate enzyme immunoassay (EIA) kit (Cozart Forensic Microplate EIA for cotinine, product no. M155B1) from Concateno (Abingdon, UK). Further details are as outlined in **Chapter 3** (**section 3.2.3.2**).

4.2.4 Long and Short Term Exposure Attributions

Annual NOx, NO_2, PM_{10} and $PM_{2.5}$ concentrations were derived using the KCL urban model (Kelly et al. 2011; Beevers et al. 2013) using ADMS dispersion model v4 and road source model v2.3 (CERC19), measured hourly meteorological data, empirically derived $NO\text{-}NO_2\text{-}O_3$ and PM relationships and emissions from the London Atmospheric Emissions Inventory (GLA 2008). Full details of the model and the selection of the 20m buffer zone around each volunteer's residence were provided in **Chapter 3** (**section 3.2.4**).

In this analysis these long term, annual exposure estimates were also weighted for periods spent at home (H) and school (S) address points based on the following criteria: that each child spending 84.4% of their time at home and 15.6% at school, based on a 7 hour school day, for 5 days per week, 39 weeks per year. Thus each child's weighted exposure was estimated by $E_{H+S} = 0.884*E_H + 0.156*E_S$. Again, as these exposure estimates were highly correlated with the home exposure attributions only the later were employed in the analysis of the final data set (**Appendix I, Figures I1-I4**).

Acute exposure estimates were derived at the address point by scaling annual mean concentrations according to a 'Nowcast' factor calculated for each pollutant for periods immediately prior to lung function evaluation. This factor (f) is defined as the ratio between concentrations measured by a local subset of continuous air pollution monitoring sites (L) in the prior period (t), and the annual mean (a) measured by the same sites. Thus, the acute exposure concentration estimate [P] at time t for point (x,y) was calculated as:

$$[P]\_t\hat{}((x,y)) = f.[P]\_a\hat{}((x,y)) \quad \text{where } f = ([P]\_t\hat{}L)/([P]\_a\hat{}L)$$

For this study 'Nowcast' scaling factors were calculated for the 24 hours and 7 days prior to the school visits, working back from 10 am on the visit day to reflect both acute and subchronic exposure periods. To derive NOx and NO_2, scaling factors measurements were averaged across 14-17 urban background and roadside sites within and surrounding the London boroughs of Tower Hamlets and Hackney, based on data availability. The decision to use both site types in deriving this factor, as opposed to the background sites only was based on the close residential proximity of many of the children within this study to busy roads. For the PM_{10} and $PM_{2.5}$ scaling factors, measurements from 9-13 and 14-20 background and roadside sites were averaged.

4.2.5 Sample Collection and SNP Genotyping

Saliva was collected from the children using the Oragene DNA kit OG-250 (DNA Genotek Inc, Canada), with genomic DNA recovered using salt out method, followed by whole genome amplification using the GenomePlex Complete Whole Genome Amplification (WGA2) Kit (Sigma Aldrich) as previously described (Arneson et al. 2008). Two panels of SNPs were investigated using the GoldenGate genotyping assay on Illumina BeadXpress platform (Illumina Inc., San Diego, USA). The confirmatory panel consisted of 7 SNPs selected as known, or hypothesized susceptibility loci including GSTM1 (rs366631), GSTP1 (rs749174, rs1695) and NQO1 (rs2917666, rs689453, rs1800566, rs10517) genes. These candidate SNPs were selected based on a review of the published literature, presented in **Chapter 1**, **Tables 1.3** and **1.4**. The second exploratory panel consisted of 9 SNPs, including 5 SNPs of AhR (rs2074113, rs2066853, rs17722841, rs17779352 and rs2282885) and 4 SNPs within the CYP1A1 gene (rs2606345, rs1799814, rs17861115 and rs2198843). For AhR, the non-synonymous SNP rs2066853 (Arg554Lys) within the transactivation domain (exon 10), which plays an essential role in regulating expression of target genes; and the synonymous SNP rs17779352 (Asn44Asn), within the basic helix-loop-helix domain (exon 2), involved in DNA biding, were selected. The former (rs2066853) represents the most widely studied AhR polymorphism, with evidence that it influences inhaled PAH metabolism in occupational settings (Chen et al., 2007; Chen et al., 2009). In vitro studies have also shown it exerts a strong influence on receptor function (Harper et al., 2002). In addition, the following Tagging SNPs were also examined, rs2074113 (33G/T); rs17722841, (4640G/A) and rs2282885, 3946G/A, all located in intervening sequences, to capture information on other SNPs not explicitly included in the present study (Ng et al., 2009).

For CYP1A1 4 SNPS were examined, one non-synonymous SNP (rs1799814, exon 7, Thr461Asn), and three Tagging SNPs (rs2606345 (606T/G), rs17861115 (-9893G/A) and rs2198843 (11599 C/G)), previously shown to be associated with markers of PAH metabolism: PAH-DNA adducts in blood (Wang et al., 2008; Abnet et al., 2007), increased

cancer risk in smokers (Rotunno et al., 2009), or increased risk of adult respiratory distress syndrome in patients with nosocomial pneumonia (Salnikova et al., 2013). The non-synonymous SNP rs1799814 (Thr461Asn) has been related to greater enzyme catalytic efficiency, related to structural changes in the haeme-binding domain.

Genotype data was analyzed for quality control using the BeadStudio software. To further improve the genotype dataset, manual editing was also performed as an additional quality control and any poor quality samples were excluded from the final analysis at this step. Hardy-Weinberg equilibrium values and genotype & allele frequencies for all SNPs are shown in **Tables I1** (**Panel 1**) and **Table I2** (**Panel 2**) in the **Appendix I**. All SNPs included for analysis were in HWE ($p>0.05$) except for SNP rs1799814 in white population and had a genotyping success rate of 99%. All SNPs displayed a >5% minor allele frequency, determined from the published web based data sets (1000 genome data set: integrated phase 1, version 3, March 2012; HapMap (Thorisson et al., 2005)). Details of linkage disequilibrium are presented in **Appendix I**, **Table I3**.

4.2.6 Statistical Analysis

Before applying any models, the distributional form of the outcomes was assessed using quantile plots of the variable against the normal distribution. Any outliers were identified using box plots and their influence on the model was evaluated using deviance residuals where possible. For each model, residuals were assessed for normality using quantile plots of the variable against the normal distribution and checked for hetero scedasticity by plotting residuals against the fitted values on the school level. For lung function, the data were normally distributed.

Linear mixed models with a random effect for school were used to examine the effects of individual air pollutant exposures on lung function outcomes. To address the influence of bias introduced by the study population, a number of baseline characteristics

were selected a priori to be included in the models: age, sex, height, body mass index (BMI), self-reported ethnicity (Asian, Black, White, Other/Mixed) and exposure to tobacco as measured by a urinary cotinine to creatinine ratio >30ng/ml. To assess the assumption of no change over time study year was also included as a covariate in the models.

For the selected SNPs, genotype was recorded as a nominal categorical variable representing wild type, heterozygous and homozygous mutant. Hardy Weinberg Equilibrium was assessed for each SNP and the Fst statistic was calculated for each SNP over reported ethnicity to assess if results from children with different ethnicities could be analysed together. In the next step, we assessed whether genotype modified the effect of air pollution on the measurements of lung function. In order to distinguish any direct effects of genotype on outcomes a crude and an adjusted analysis was first performed for each SNP. The adjusted analysis included the same covariates as described above but no air pollution exposures. Next, to test for effect modification of the SNPs, interaction terms of each categorical SNP and each continuous air pollution variable were added to the linear mixed models. For each SNP the wild type genotype was used as the reference category for the comparison between the three groups. For cases where the effect modification was significant for only one comparison but not the other a global Wald test was performed to assess influence of the SNP overall. Additionally we performed all analysis combining heterozygous and homozygous mutant into one category, which increased the power to detect effects of mutations in each SNP and also aimed to confirm results from the global tests. This method has been applied in other gene modification studies.

Multiple testing was performed on all significant and non-significant p-values using a step-up false discovery rate (FDR) procedure as originally proposed by Simes (1986) and later justified by Benjamini and Yekutieli (2001). As described by both groups controlling the false discovery rate is less conservative than using the family-wise error rate and therefore results in more differences being significant at the cost of some confidence in the results. The Simes FDR procedure calculates a corrected critical p-value for each p-value of interest (effects of air pollutant variables or interaction terms), so that an individual null

hypothesis can be rejected if its corresponding p-value is smaller than the corrected critical p-value.

Models were examined on the 5% significance level for two-sided tests of individual associations and on the 10% significance level for interaction terms of effect modification. All statistical analyses were performed using Stata 10.1 (StataCorp, College Station, TX, USA).

4.3 RESULTS

4.3.1 Subject Demographics

Details of participant characteristics are presented in **Table 4.1**, for each individual year and combined across the first three years of the study. The majority of the study population was of central Asian background, predominately Bangladeshi. Data on lung function were available for 1,000 children at an average age of 8.8 years. Of these, technically acceptable FEV_1 and FVC measures were available for 951 and 920 children respectively. The smoking exposure status (urinary cotinine) was obtained for 954 children; of whom 222 had an exposure level above 25 ng/ml and were therefore considered as being exposed to environmental tobacco smoke. Following adjustment for sample dilution using urinary creatinine 212 children were found to be above the threshold. **Table 4.2** illustrates the post-bronchodilator FVC and FEV_1 measurements for all of the subjects from whom complete sets of covariant were available and were thus included in the full analysis. The mean FEV_1 and FVC combined across the first 3-years of the study (2008-11) were 1,692 ml and 1,891 ml, respectively. Saliva samples were collected for 182 children from 10 schools (Nov – Feb 2008/2009), 415 children from 19 schools (Nov/Feb 2009/2010) and 404 children from 22 schools (Nov – Mac 2010/2011). Of the 1001 samples collected from children during the school visits for Year 1-3, 13 (1.3%) samples were lost due to technical

error during samples handling. DNA was successfully extracted from all of the remaining 988 saliva samples, but subsequently 14 poorly performing samples were excluded to improve the genotype data set leaving 974 samples in total for downstream analysis. Of the 974 children with genotype data, 894 performed acceptable FEV_1 (92%), and 864 (89%) had a technically acceptable FVC measurement.

4.3.2 Exposure Attributions

Whilst the spatial resolution of the dispersion model was good I also explored whether short term exposure might influence the lung function determinations on the day of the actual school visit, or in more extended periods running up to the visit date. The importance of considering visit date is illustrated in **Figure 4.1**, which shows the daily variation in NO_2 concentrations at roadside and urban background sites surrounding the study area over the three successive winter sampling periods. The distance of the school from a major road was also illustrated. Addressing the influence of acute air pollution exposures on lung function is a critical question as it has been shown that exposure to both diesel (Holgate et al. 2003a, 2003b) and roadside (McCreanor et al. 2007) air pollution can elicit transient decrements in lung function. To date only two studies have attempted to address this issue, by including daily background measurements from their study area as a covariant in their regressions model (Oftedal et al. 2008; Schultz et al. 2012). Whilst this approach has some merit, it does not address acute versus chronic pollutant exposures on the same geographical scale, which is problematic in urban settings where exposures are highly influenced by a subject's residential proximity to roads. A straightforward method of aiding the demarcation of acute effects, [triggered by recent pollutants exposures] from chronic effects [related to longer term exposures] was developed for this study.

Table 4.1 Descriptive data for study participants by year and for year 1-3 combined (for all those who attempted spirometry)

Variables	**Yr 1**	**Yr 2**	**Yr 3**	**Yrs 1-3**
	(n = 182)	(n = 414)	(n = 404)	(n = 1000)
Girls, n (%)	77 (42.3)	198 (47.8)	208 (51.5)	483 (48.3)
Age, yrs; mean (SD)	8.8 (0.3)	8.8 (0.3)	8.8 (0.3)	8.8 (0.3)
Height, cm; mean (SD)	133.9 (6.4)	133.8 (7.0)	133.7 (6.6)	133.8 (6.7)
Weight, kg; mean (SD)	32.5 (7.6)	32.8 (8.3)	32.4 (7.8)	32.6 (7.9)
BMI, kg/m^2; mean (SD)	17.9 (3.0)	18.1 (3.3)	17.9 (3.4)	18.0 (3.3)
Ethnicity, n (%)				
- Asian	69 (37.9)	152 (36.7)	138 (34.2)	359 (35.9)
- Black	45 (24.7)	104 (25.1)	101 (25.0)	250 (25.0)
- White	53 (29.1)	115 (27.8)	104 (25.7)	272 (27.2)
- Other	15 (8.2)	43 (10.4)	61 (15.1)	119 (11.9)
IMD score; mean (SD)	46.0 (11.6)	45.6 (10.1)	44.8 (12.0)	45.3 (11.2)
ETS exposure, n (%)	51 (28.0)	95 (22.9)	66 (16.3)	212 (21.2)

MD score = IMD 2010; ETS exposure = cotinine: creatinine ratio >30 ng/mg. Percentages may not sum exactly to 100 % due to rounding.

Notes: Of the 359 subjects coded as Asian, 87% (314) gave their self reported ethnicity as Bangladeshi. The population, reporting themselves as black, was more diverse, with the major contributions being Caribbean (19%), Nigerian (18%), other black (27%), or black African (14%). The mixed other group consisted of 46% of mixed and 54% other, where it was difficult to discriminate continental origin, with the largest group being children reporting themselves as Turkish (26%).

Table 4.2 Summary lung function data, restricted to subjects with complete sets of co-variants

Variable	N	Mean	SD	Range	Median
FEV_1, L	951	1.69	0.28	0.84 - 2.73	1.68
FEV_1, % predicted	950	94.4	11.8	59.4 - 131.3	94.3
FVC, L	920	1.89	0.32	1.11 - 3.19	1.87
FVC, % predicted	919	92.3	12.1	58.2 - 128.9	92.1

Chronic exposures were assessed using annual mean concentration maps at 20m resolution, with subjects assigned exposure levels based on their residential and/or school address. Acute exposures were then assessed by scaling these annual mean concentrations according to a 'Nowcast' factor calculated for each pollutant for the period immediately prior to the biomarker extraction. This factor was defined as the ratio between concentrations measured by a local subset of London Air Quality Network monitoring sites in the prior period, and the annual mean measured by the same sites. Nowcast factors varied between 0.4 and 3.0 for $PM_{2.5}$, i.e., $PM_{2.5}$ concentrations assigned to the subject's address for the acute assessment were up to three times higher than the chronic assessment. This temporal variation was far greater than the annual mean spatial variation in $PM_{2.5}$. A similar pattern was apparent for each of the modelled pollutant species. Using these approaches exposure attribution for each individual were derived based on annual (chronic exposure), 7 day (referred to throughout as sub-chronic) and 24 hour (acute exposures). Descriptive data on each of the modelled pollutant data are presented in **Table 4.3** for the 1,000 children who attempted spirometry, based on home residential address and weighted for periods spent at the school location. This approach was based on observations made by McConnell et al (2010) where combined school and home exposure was related to childhood incident asthma.

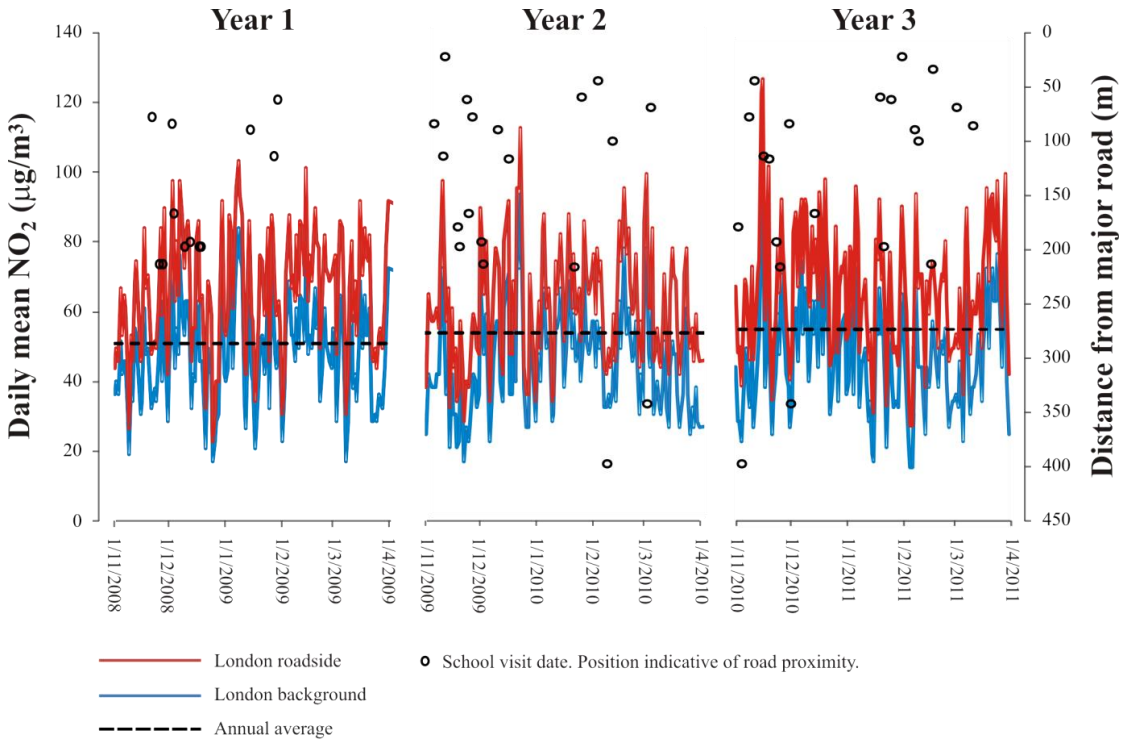


Figure 4.1 Daily mean NO_2 variation at roadside (red line) and urban back ground (blue line) sites within and bordering the London boroughs of Tower Hamlets and Hackney over the first three years of the LEZ study. The dates of the school visits are illustrated with open circles with the position on the y-axis reflecting the proximity of the school to major roads. The dashed line represents the mean London NO_2 concentration for each study period across all LAQN monitoring sites.

From the data presented in **Table 4.3** it was apparent that there was little difference between the exposure estimates based on residential address alone, or weighted for the time spent in school, probably reflecting the close proximity of the children to their primary schools in Central London. The 'Home' and 'Average' exposure attributions were also highly correlated (**Appendix I, Figures I1 – I4**), therefore the decision was made to focus only on the former, to limit the number of exposure metrics applied to the health data. It is also important to note that each of the four pollutants modelled at the same temporal scale

were highly correlated (**Table 4.4**), which precluded the subsequent use of multi-pollutant models in the subsequent analysis. The degree of association between the pollutants significantly degraded as shorter exposure intervals were considered.

Table 4.3 Descriptive exposure data on the 1000 participants attempting lung function testing across the first three years of the Low Emission Zone Study. Exposures are expressed as annual, weekly, or 24 hour average ($\mu g/m^3$) exposures at the residential address level

Pollutant	**N**	**Mean**	**SD**	**Min**	**Max**	**Median**
Home NOx *(Annual)*	996	75.74	14.62	47.52	235.44	73.11
Home + school NOx *(Annual)*	996	78.83	13.86	53.06	224.17	76.39
Home NOx *(7 day)*	996	46.51	16.83	16.90	181.29	45.93
Home + school NOx *(7 day)*	996	48.39	16.93	17.93	172.61	48.67
Home NOx *(24 hours)*	996	58.14	43.40	9.49	396.17	46.35
Home + school NOx *(24 hours)*	996	60.42	44.40	9.97	389.04	48.67
Home NO_2 *(Annual)*	996	43.52	5.45	31.16	98.93	42.51
Home + school NO_2 *(Annual)*	996	45.30	5.20	34.05	95.89	44.38
Home NO_2 *(7 day)*	996	24.99	6.04	12.21	67.27	25.24
Home + school NO_2 *(7 day)*	996	26.01	6.13	12.83	65.21	26.42
Home NO_2 *(24 hours)*	996	27.76	10.90	8.12	75.45	27.35
Home + school NO_2 *(24 hours)*	996	28.88	11.16	8.50	75.20	28.63
Home PM_{10} *(Annual)*	996	23.38	1.44	20.41	33.16	23.06
Home + school PM_{10} *(Annual)*	996	24.30	1.35	21.60	34.00	24.06
Home PM_{10} *(7 day)*	996	22.69	5.49	13.55	60.00	22.35
Home + school PM_{10} *(7 day)*	996	23.60	5.73	14.10	61.03	23.31
Home PM_{10} *(24 hours)*	996	24.30	9.37	8.67	65.98	23.62
Home + school PM_{10} *(24 hours)*	996	25.27	9.73	9.02	66.77	24.65
Home $PM_{2.5}$ *(Annual)*	996	13.73	0.82	11.68	17.97	13.62
Home + school $PM_{2.5}$ *(Annual)*	996	14.27	0.81	12.38	18.27	14.17
Home $PM_{2.5}$ *(7 day)*	996	13.65	3.71	8.16	35.30	13.16
Home + school $PM_{2.5}$ *(7 day)*	996	14.20	3.87	8.49	36.03	13.71
Home $PM_{2.5}$ *(24 hours)*	996	14.89	6.50	6.32	44.95	13.70
Home + school $PM_{2.5}$ *(24 hours)*	996	15.49	6.76	6.61	45.88	14.25

Table 4.4 Associations (Pearson correlations) between annual, 7-day and 24 hour exposure attributions

		Annual				7 day				24 hours			
		NOx	NO_2	PM_{10}	$PM_{2.5}$	NOx	NO_2	PM_{10}	$PM_{2.5}$	NOx	NO_2	PM_{10}	$PM_{2.5}$
Annual	NOx	1.00											
	NO_2	1.00	1.00										
	PM_{10}	0.93	0.94	1.00									
	$PM_{2.5}$	0.77	0.78	0.87	1.00								
7 day	NOx	0.39	0.39	0.33	0.10	1.00							
	NO_2	0.41	0.42	0.34	0.12	0.95	1.00						
	PM_{10}	0.26	0.26	0.20	0.07	0.51	0.46	1.00					
	$PM_{2.5}$	0.15	0.15	0.11	0.02	0.52	0.45	0.95	1.00				
24 hours	NOx	0.21	0.22	0.24	0.15	0.56	0.49	0.11	0.15	1.00			
	NO_2	0.29	0.30	0.30	0.18	0.59	0.59	0.10	0.11	0.91	1.00		
	PM_{10}	0.19	0.20	0.22	0.18	0.42	0.38	0.25	0.23	0.80	0.76	1.00	
	$PM_{2.5}$	0.11	0.12	0.13	0.11	0.41	0.36	0.21	0.25	0.79	0.72	0.94	1.00

4.3.3 Lung Function Associations with Modeled Exposures

The adjusted coefficients for the association of the demographic variables with lung function, adjusted for all pollutant attributions are summarized in **Table 4.5**. For both FEV_1 and FVC significant associations were observed with sex, height and BMI, but not IMD score or ETS exposure, confirmed using the urinary cotinine to creatinine ratio. Clear differences in lung function variables between the ethnicities were observed with the Black and Asian children having smaller lung volumes in accord with recent reports (Kirkby et al. 2013). No impact of study year was seen on the measurements obtained.

Table 4.5 Adjusted coefficients for the association between demographic variables and measures of lung function

Outcome	DISPERSION MODELLING	NOWCAST MODELLING 7 DAYS	NOWCAST MODELLING 1 DAY
FEV1	**Annual**	**7-days**	**1-day**
Age (years)	0.0165[-0.0233,0.0563]	0.0202[-0.0204,0.0608]	0.0186[-0.0212,0.0585]
Sex (Girls vs Boys)	-0.0821***[-0.1069,-0.0573]	-0.0840***[-0.1088,-0.0592]	-0.0822***[-0.1070,-0.0574]
Height (cm)	0.0259***[0.0237,0.0280]	0.0259***[0.0237,0.0281]	0.0259***[0.0237,0.0281]
BMI	0.0099***[0.0059,0.0139]	0.0096***[0.0056,0.0136]	0.0099***[0.0058,0.0139]
IMD Score	-0.0007[-0.0019,0.0004]	-0.0005[-0.0016,0.0006]	-0.0005[-0.0016,0.0007]
ETS Score	0.0028[-0.0287,0.0343]	0.0002[-0.0311,0.0315]	0.0019[-0.0295,0.0334]
Ethnicity Black vs Asian	-0.1315***[-0.1656,-0.0975]	-0.1343***[-0.1685,-0.1002]	-0.1347***[-0.1691,-0.1003]
Ethnicity White vs Asian	0.1099***[0.0767,0.1432]	0.1040***[0.0703,0.1377]	0.1051***[0.0711,0.1391]
Ethnicity Others vs Asian	0.1090***[0.0676,0.1503]	0.1035***[0.0617,0.1452]	0.1062***[0.0645,0.1478]
Study Year 2 vs 1	-0.0356[-0.1610,0.0898]	-0.0129[-0.0533,0.0275]	-0.0199[-0.0601,0.0204]
Study Year 3 vs 1	-0.0072[-0.0527,0.0383]	0.0162[-0.0261,0.0585]	0.002[-0.0388,0.0429]
NOx (ug/m^3)	-0.0084[-0.0176,0.0008]	-0.001[-0.0036,0.0017]	0[-0.0008,0.0008]
NO_2 (ug/m^3)	0.0236[-0.0032,0.0503]	0.0057[-0.0015,0.0128]	0.0015[-0.0017,0.0046]
PM_{10} (ug/m^3)	-0.0127[-0.0712,0.0459]	-0.0067[-0.0144,0.0010]	-0.0022[-0.0067,0.0023]
$PM_{2.5}$ (ug/m^3)	0.0126[-0.1097,0.1349]	0.0068[-0.0044,0.0179]	0.0018[-0.0042,0.0078]
FVC			
Age (years)	0.0291[-0.0170,0.0753]	0.0295[-0.0175,0.0766]	0.0317[-0.0146,0.0780]
Sex (Girls vs Boys)	-0.1424***[-0.1711,-0.1138]	-0.1433***[-0.1720,-0.1146]	-0.1421***[-0.1709,-0.1134]
Height (cm)	0.0280***[0.0254,0.0305]	0.0279***[0.0253,0.0304]	0.0278***[0.0252,0.0303]
BMI	0.0153***[0.0106,0.0199]	0.0148***[0.0101,0.0194]	0.0151***[0.0104,0.0197]
IMD Score	-0.0004[-0.0017,0.0010]	-0.0001[-0.0014,0.0012]	-0.0003[-0.0016,0.0011]
ETS Score	0.0035[-0.0327,0.0396]	0.0029[-0.0331,0.0389]	0.0033[-0.0328,0.0395]
Ethnicity Black vs Asian	-0.1312***[-0.1713,-0.0912]	-0.1341***[-0.1735,-0.0948]	-0.1290***[-0.1694,-0.0886]
Ethnicity White vs Asian	0.1524***[0.1133,0.1915]	0.1514***[0.1126,0.1902]	0.1529***[0.1131,0.1928]
Ethnicity Others vs Asian	0.1141***[0.0657,0.1625]	0.1111***[0.0626,0.1595]	0.1177***[0.0690,0.1664]
Study Year 2 vs 1	-0.0548[-0.1985,0.0889]	-0.0457[-0.0922,0.0008]	-0.0509*[-0.0979,-0.0038]
Study Year 3 vs 1	0.0078[-0.0447,0.0603]	0.0201[-0.0285,0.0686]	0.0064[-0.0412,0.0541]
NOx (ug/m^3)	-0.0067[-0.0173,0.0039]	-0.0029[-0.0060,0.0002]	0.0002[-0.0008,0.0012]
NO_2 (ug/m^3)	0.0147[-0.0161,0.0456]	0.0087*[0.0005,0.0170]	-0.0001[-0.0038,0.0036]
PM_{10} (ug/m^3)	-0.0039[-0.0713,0.0635]	-0.0048[-0.0136,0.0040]	0.0013[-0.0041,0.0066]
$PM_{2.5}$ (ug/m^3)	0.0124[-0.1276,0.1524]	0.0056[-0.0073,0.0186]	-0.0032[-0.0102,0.0038]

The relationship between post-bronchodilator FVC and the modeled pollutant concentrations, across the acute to chronic exposure windows are illustrated in **Figure 4.2**. Data are presented separately for each individual year and aggregated across the first three years of the study for additional power. Overall no evidence of any association between FVC and the acute (24h) or sub-chronic (7 day) exposure attributions were observed, and there was no clear pattern in the year by year data consistent with a beneficial effect of the LEZ over the first three years of its lifetime. Significant associations ($P<0.05$) were apparent at the nominal level for all four pollutants in annual exposure attributions with FVC, with the association with NOx remaining significant after correction for multiple testing to account for the different exposure periods examined. In contrast, no associations

were noted between any of the exposure attributions examined and FEV_1 (**Table 4.6a** and **4.6b**). As a secondary analysis, we also explored how these annual associations with lung capacity varied across the 4 ethnic groups. This analysis, illustrated in **Figure 4.3** demonstrated that the reduction in FVC associated with annual NOx and NO_2 exposures was predominately experienced in the white population,.

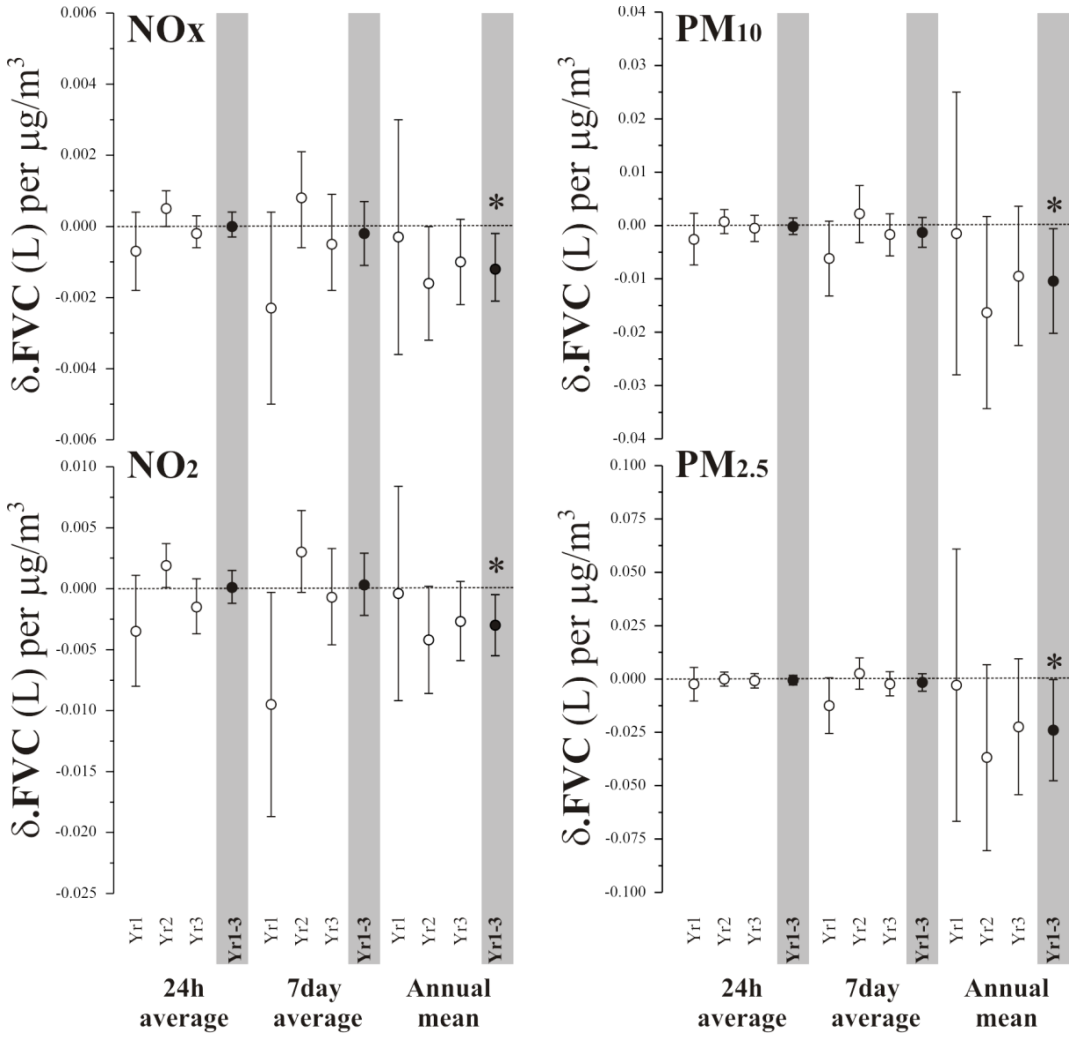


Figure 4.2 Association between each of the 4 selected pollutant attributions, based on 24 hour, 7 day and annual exposures and the measured FVC, post bronchodilator. Effect estimates are presented with 95% confidence intervals, based on exposures model by home residential address only (filled circles), and weighted for periods spent at school (open circles). Data are presented by year, as well as pooled across the first three years of the study: * $p<0.05$.

Table 4.6a Effect estimates for the association between annual NOx and NO_2 exposures with lung function. Data are expressed per $\mu g/m^3$ of pollutant and per interquartile range (IQR) of exposure

		NOx (20m)			**NO_2 (20m)**		
		b	95th CI	IQR	b	95th CI	IQR
Post FEV_1 ($L.\mu g.m^{-3}$)	24 hour	0.0001	[-0.0002,0.0004]	0.004	0.0006	[-0.0005,0.0018]	0.008
	7 day	0.0006	[-0.0002,0.0014]	0.014	0.002	[-0.0002,0.0042]	0.017
	Annual	-0.0004	[-0.0012,0.0004]	-0.004	-0.0008	[-0.0030,0.0014]	-0.003
Post FVC ($L.\mu g.m^{-3}$)	24 hour	0.0000	[-0.0003,0.0004]	0.000	0.0001	[-0.0012,0.0015]	0.001
	7 day	-0.0002	[-0.0011,0.0007]	-0.005	0.0003	[-0.0022,0.0029]	-0.041
	Annual	-0.0012*	[-0.0021,-0.0002]	-0.012*	-0.0030*	[-0.0055,-0.0005]	-0.012*

Table 4.6b Effect estimates for the association between PM_{10} and $PM_{2.5}$ with lung function. Data are expressed per $\mu g/m^3$ of pollutant and per interquartile range (IQR) of exposure

		PM_{10} (20m)			**$PM_{2.5}$ (20m)**		
		b	95th CI	IQR	b	95th CI	IQR
Post FEV_1	24 hour	0.0001	[-0.0012,0.0015]	0.001	0.0003	[-0.0016,0.0022]	0.002
	7 day	-0.0007	[-0.0031,0.0017]	-0.004	-0.0004	[-0.0038,0.0031]	-0.002
	Annual	-0.0032	[-0.0118,0.0053]	-0.004	-0.007	[-0.0276,0.0137]	-0.009
Post FVC	24 hour	-0.0002	[-0.0017,0.0014]	-0.003	-0.0006	[-0.0028,0.0017]	-0.023
	7 day	-0.0013	[-0.0041,0.0015]	-0.007	-0.0016	[-0.0057,0.0025]	-0.008
	Annual	-0.0104*	[-0.0202,-0.0006]	-0.014*	-0.0240*	[-0.0477,-0.0003]	-0.030*

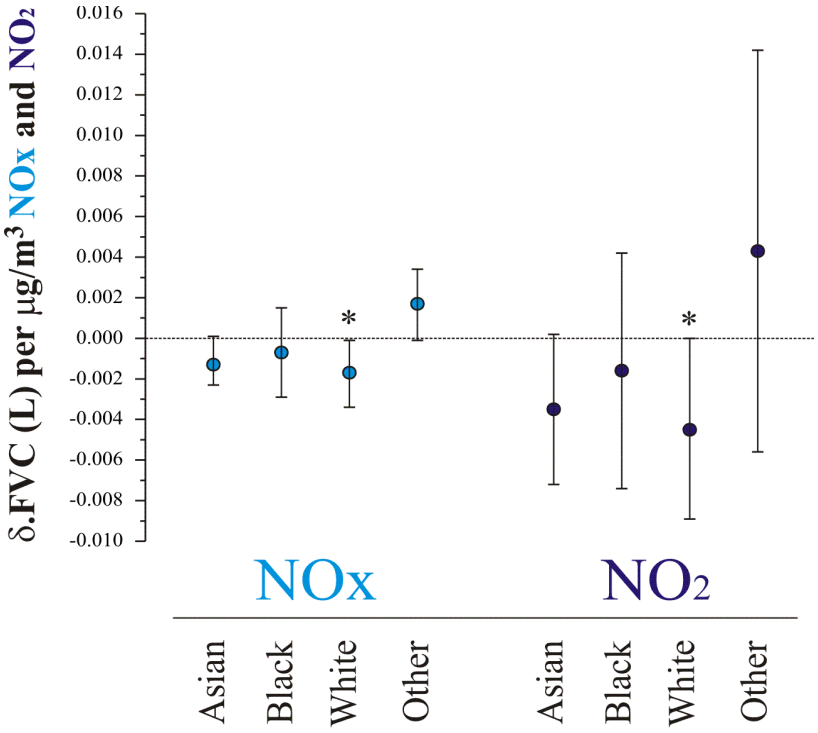


Figure 4.3 The impact of ethnicity on the associations between annual NOx and NO_2 exposure attributions with FVC.

4.3.4 Impact of SNPs in Phase II Xenobiotic Genes on Air Pollution, Lung Function Interactions - Confirmatory Analysis

The distribution of the genotype and allele frequencies for GSTP1, GSTM1 and NQO1 SNPs for total population and for the population stratified by ethnicity is shown in **Table 4.7**. Overall, all the four ethnic groups in the study population, shared similar common genotypes except for GSTM1 SNP rs366631 and NQO1 SNP rs2917666. Pseudo-SNP GSTM1 rs366631 (C/T), was used to infer GSTM1 deletion based on methods described by Huang et al. (2009). Hence, it only distinguishes whether there is a homozygous deletion (null) or not, and is unable to differentiate between the numbers of alternative alleles. Therefore, when both copies of the upstream region were deleted, the

genotype call was made solely based on homozygous variants (TT). Approximately, 45% of the Asian and 'Others' exhibited the GSTM1 deletion. In contrast, Black and White demonstrated ~30% and ~60% of GSTM deletion respectively. These values were within the ranges previously reported for these ethnic groups (Carlsten et al. 2008). With regards to GSTP1 SNP rs749174 (C/T), all ethnicities shared the same common genotype 'CC' with MAF varying between 22% to 36%, with the White population having the highest MAF and the Asian the lowest. For SNP rs1695 (Ile/Val), 'Ile/Ile' was the common genotype for all ethnicities, with the MAF showing little variation between the Black (45%) and other ethnic groups (ranging between 25% and 36%). All NQO1 SNPs except rs2917666 showed same common genotype for all the ethnic groups. Whilst Asian, Black and 'Others' exhibited 'G' as the minor allele for SNP rs2917666 with a MAF between 41% to 48%, in the White population 'C' was the minor allele with a MAF of 31%. With respect to SNPs rs689453 and rs10517, all ethnic groups had a MAF of approximately 4% and 18% respectively. In addition, despite sharing the same common genotype in SNP rs1800566 (Pro/Ser) with the rest of ethnic groups, the Asian populations had the highest MAF of the allele Ser with ~34% compared to Black who had the lowest MAF with ~17%.

Tables I3 in **Appendix I** shows the crude and adjusted effects of the studied GSTs and NQO1 genotypes on FEV_1 and FVC. Following model adjustment for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year,GSTP1 (rs1695, Ile105Val) was found to be associated with significant reductions in both FEV_1 (-32mL) and FVC (-42mL) when Ile/Val+Val/Val genotypes were combined and compared against the reference genotype (Ile/Ile). In contrast, NQO1 (rs1800566, Pro187Ser) was associated with a significant increase in FVC (+33mL) when the Pro/Ser + Ser/Ser genotype were combined and compared against Pro/Pro. Despite these associations with basal lung function I observed no significant modifying effects of polymorphisms in the studied GSTs or NQO1 genes on the relationship between lung function and annual (FEV_1, **Table 4.8**; FVC, **Table 4.9**), sub-chronic (**Appendix I**: FEV_1, **Table I6**; FVC, **Figure I7**), or acute (**Appendix I**: FEV_1, **Table I4**; FVC, **Figure I5**) exposures.

Table 4.7 Distribution of genotype and allele frequencies in GSTM1, GSTP1 and NQO1

Genotype/Allele Frequencies	Ethnicity								Total Population	
	Asian		Black		White		Others			
	n	%	n	%	n	%	n	%	n	%
GSTM1										
rs366631 C/T										
CC	0	0	0	0	0	0	0	0	0	0.0
CT	195	56.4	171	69.2	112	42.1	60	52.6	481	55.8
TT	151	43.6	76	30.8	154	57.9	54	47.4	381	44.2
Allele C	195	28.2	171	34.6	112	21.1	60	26.3	481	27.9
Allele T	497	71.8	323	65.4	420	78.9	168	73.7	1243	72.1
GSTP1										
rs749174 Intron 5 C/T										
CC	213	61.4	142	57.5	108	40.6	56	49.1	453	52.5
CT	113	32.6	90	36.4	125	47.0	48	42.1	339	39.3
TT	21	6.1	15	6.1	33	12.4	10	8.8	71	8.2
Allele C	539	77.7	374	75.7	341	64.1	160	70.2	1245	72.1
Allele T	155	22.3	120	24.3	191	35.9	68	29.8	481	27.9
rs1695 Exon 5 Ile105Val										
Ile/Ile	191	55.0	82	33.2	110	41.5	49	43.0	377	43.7
Ile/Val	132	38.0	110	44.5	118	44.5	51	44.7	370	42.9
Val/Val	24	6.9	55	22.3	37	14.0	14	12.3	115	13.3
Alelle Ile	514	74.1	274	55.5	338	63.8	149	65.4	1124	65.2
Alelle Val	180	25.9	220	44.5	192	36.2	79	34.6	600	34.8
NQO1										
rs2917666 3´ UTR (C/G)										
CC	76	21.9	39	15.8	123	46.2	24	21.1	228	26.4
CG	163	47.0	123	49.8	120	45.1	62	54.4	424	49.1
GG	108	31.1	85	34.4	23	8.6	28	24.6	211	24.4
Allele C	315	45.4	201	40.7	366	68.8	110	48.2	880	51.0
Allele G	379	54.6	293	59.3	166	31.2	118	51.8	846	49.0
rs689453 Exon 2 Glu24 (G/A)										
GG	336	96.8	221	89.5	238	89.5	101	88.6	796	92.2
GA	11	3.2	25	10.1	27	10.2	13	11.4	65	7.5
AA	0	0.0	1	0.4	1	0.4	0	0.0	2	0.2
Allele G	683	98.4	467	94.5	503	94.5	215	94.3	1657	96.0
Alelle A	11	1.6	27	5.5	29	5.5	13	5.7	69	4.0
rs1800566 Exon 6 Pro187Ser										
Pro/Pro	153	44.1	173	70.0	179	67.3	58	50.9	502	58.2
Pro/Ser	154	44.4	66	26.7	78	29.3	50	43.9	311	36.0
Ser/Ser	40	11.5	8	3.2	9	3.4	6	5.3	50	5.8
Allele Pro	460	66.3	412	83.4	436	82.0	166	72.8	1315	76.2
Alelle Ser	234	33.7	82	16.6	96	18.0	62	27.2	411	23.8
rs10517 3´ UTR region (C/T)										
CC	220	63.4	171	69.5	197	74.1	70	61.4	581	67.4
CT	116	33.4	68	27.6	67	25.2	39	34.2	257	29.8
TT	11	3.2	7	2.8	2	0.8	5	4.4	24	2.8
Allele C	556	80.1	410	83.3	461	86.7	179	78.5	1419	82.3
Alelle T	138	19.9	82	16.7	71	13.3	49	21.5	305	17.7

Minor allele frequency for this ethnicity is different with other ethnicities for the particular SNP.
Genotype distribution for this ethnicity are not similar with other ethnicity groups despite sharing same major/minor allele.
Similar genotype frequencies for both ethnicity group for the particular SNP.

Allele frequencies calculation for wild type (AA) = AA/(AA+AB+BB)
Allele frequencies calculation for heterozygotes (AB) = AB/(AA+AB+BB)
Allele frequencies calculation for homozygotes (BB) = BB/(AA+AB+BB)

Table 4.8 Effects of GSTM1, GSTP1 and NQO1 genotypes on FEV_1 outcomes by annual mean pollutant exposure

		FEV_1			
Genotype	**Subjects**	**NOx β (95% CI)**	**NO_2 β (95% CI)**	**PM_{10} β (95% CI)**	**$PM_{2.5}$ β (95% CI)**
GSTM1					
rs366631 (C/T)					
CT	482	Reference	Reference	Reference	Reference
TT	383	0.0007[-0.0012,0.0025]	0.0016[-0.0032,0.0064]	0.0051[-0.0122,0.0225]	0.0193[-0.0111,0.0496]
GSTP1					
rs749174 (intron 5 C/T)					
CC	455	Reference	Reference	Reference	Reference
CT	340	0.0006[-0.0010,0.0023]	0.0016[-0.0029,0.0060]	0.0066[-0.0105,0.0236]	0.0180[-0.0131,0.0491]
TT	71	0.0036#[-0.0007,0.0079]	0.0087[-0.0022,0.0197]	0.0170[-0.0223,0.0562]	0.0210[-0.0367,0.0788]
CT + TT	411	0.0008[-0.0008,0.0024]	0.0021[-0.0022,0.0064]	0.0072[-0.0093,0.0238]	0.0181[-0.0113,0.0476]
rs1695 (Exon 5 Ile105Val)					
Ile/Ile	379	Reference	Reference	Reference	Reference
Ile/Val	371	0.0004[-0.0014,0.0022]	0.0010[-0.0037,0.0058]	0.0029[-0.0149,0.0207]	-0.0083[-0.0403,0.0236]
Val/Val	115	0.0011[-0.0012,0.0033]	0.0032[-0.0031,0.0094]	0.0090[-0.0173,0.0353]	0.0098[-0.0362,0.0557]
Ile/Val + Val/Val	486	0.0006[-0.0010,0.0022]	0.0016[-0.0027,0.0060]	0.0041[-0.0126,0.0208]	-0.0042[-0.0340,0.0256]
NQO1					
rs2917666 (3´ UTR (C/G))					
CC	231	Reference	Reference	Reference	Reference
CG	424	0.0010[-0.0008,0.0029]	0.0028[-0.0021,0.0078]	0.0116[-0.0079,0.0312]	0.0197[-0.0158,0.0551]
GG	211	0.0002[-0.0019,0.0023]	0.0007[-0.0050,0.0065]	0.0070[-0.0149,0.0290]	0.0135[-0.0268,0.0539]
CG + GG	635	0.0007[-0.0009,0.0024]	0.0021[-0.0024,0.0065]	0.0099[-0.0079,0.0278]	0.0173[-0.0153,0.0499]
rs689453 (Exon 2 Glu24 (G/A))					
GG	799	Reference	Reference	Reference	Reference
GA	65	-0.0004[-0.0046,0.0039]	-0.0014[-0.0123,0.0094]	-0.0064[-0.0440,0.0311]	-0.0045[-0.0636,0.0545]
AA	2	-0.0101[-0.0347,0.0144]	-0.0264[-0.0898,0.0371]	-0.0780[-0.2669,0.1108]	-0.1947[-0.6650,0.2756]
GA + AA	67	-0.0003[-0.0044,0.0038]	-0.0013[-0.0118,0.0092]	-0.0059[-0.0418,0.0301]	-0.0019[-0.0587,0.0550]
rs1800566 (Exon 6 Pro187Ser)					
Pro/Pro	505	Reference	Reference	Reference	Reference
Pro/Ser	311	0.0004[-0.0015,0.0024]	0.0010[-0.0042,0.0061]	0.0051[-0.0136,0.0238]	0.0044[-0.0281,0.0369]
Ser/Ser	50	-0.0015[-0.0040,0.0010]	-0.0041[-0.0108,0.0026]	-0.0089[-0.0348,0.0170]	-0.0060[-0.0591,0.0472]
Pro/Ser + Ser/Ser	361	-0.0003[-0.0019,0.0014]	-0.0008[-0.0052,0.0036]	0.0006[-0.0159,0.0172]	0.0014[-0.0283,0.0312]
rs10517 (3´ UTR (C/T))					
CC	584	Reference	Reference	Reference	Reference
CT	257	0.0001[-0.0018,0.0021]	0.0005[-0.0048,0.0058]	0.0005[-0.0192,0.0201]	-0.0076[-0.0418,0.0265]
TT	24	0.0026[-0.0050,0.0102]	0.0058[-0.0131,0.0248]	0.0051[-0.0557,0.0659]	-0.0061[-0.1011,0.0888]
CT + TT	281	0.0003[-0.0016,0.0022]	0.0009[-0.0042,0.0060]	0.001[-0.0179,0.0199]	-0.0068[-0.0395,0.0259]

Table 4.9 Effects of GSTM1, GSTP1 and NQO1 genotypes on FVC outcomes by annual mean pollutant exposure

		FVC			
Genotype	**Subjects**	**NOx β (95% CI)**	**NO_2 β (95% CI)**	**PM_{10} β (95% CI)**	**$PM_{2.5}$ β(95% CI)**
GSTM1					
rs366631 (C/T)					
CT	482	Reference	Reference	Reference	Reference
TT	383	0.0000[-0.0021,0.0021]	0.0001[-0.0054,0.0057]	0.0021[-0.0179,0.0222]	0.0043[-0.0309,0.0395]
GSTP1					
rs749174 (intron 5 C/T)					
CC	455	Reference	Reference	Reference	Reference
CT	340	0.0008[-0.0011,0.0027]	0.0020[-0.0031,0.0070]	0.0070[-0.0125,0.0265]	0.0221[-0.0136,0.0577]
TT	71	0.0015[-0.0039,0.0068]	0.0030[-0.0109,0.0169]	0.0055[-0.0422,0.0532]	0.0242[-0.0444,0.0929]
CT + TT	411	0.0008[-0.0010,0.0027]	0.002[-0.0029,0.0069]	0.0067[-0.0122,0.0256]	0.0222[-0.0116,0.0561]
rs1695 (Exon 5, Ile105Val)					
Ile/Ile	379	Reference	Reference	Reference	Reference
Ile/Val	371	0.0003[-0.0017,0.0024]	0.0008[-0.0045,0.0062]	0.0019[-0.0185,0.0222]	-0.0097[-0.0464,0.0270]
Val/Val	115	0.0011[-0.0015,0.0037]	0.0029[-0.0044,0.0101]	0.0053[-0.0252,0.0357]	0.0113[-0.0421,0.0647]
Ile/Val + Val/Val	486	0.0006[-0.0013,0.0024]	0.0014[-0.0036,0.0063]	0.0025[-0.0167,0.0216]	-0.005[-0.0392,0.0293]
NQO1					
rs2917666 (3´ UTR (C/G))					
CC	231	Reference	Reference	Reference	Reference
CG	424	0.0009[-0.0011,0.0030]	0.0028[-0.0029,0.0084]	0.0105[-0.0120,0.0330]	0.0127[-0.0282,0.0536]
GG	211	0.0002[-0.0022,0.0026]	0.0011[-0.0054,0.0076]	0.0055[-0.0195,0.0305]	0.0112[-0.0349,0.0573]
CG + GG	635	0.0007[-0.0012,0.0025]	0.0022[-0.0029,0.0073]	0.0088[-0.0116,0.0292]	0.0122[-0.0253,0.0497]
rs689453 (Exon 2, Glu24 (G/A))					
GG	799	Reference	Reference	Reference	Reference
GA	65	-0.0013[-0.0061,0.0036]	-0.0036[-0.0160,0.0089]	-0.0110[-0.0541,0.0322]	-0.0214[-0.0897,0.0468]
AA	2	-0.0050[-0.0328,0.0227]	-0.0131[-0.0849,0.0588]	-0.0385[-0.2527,0.1756]	-0.0973[-0.6305,0.4360]
GA + AA	67	-0.0013[-0.0060,0.0034]	-0.0036[-0.0156,0.0084]	-0.0112[-0.0525,0.0301]	-0.021[-0.0866,0.0445]
rs1800566 (Exon 6, Pro187Ser)					
Pro/Pro	505	Reference	Reference	Reference	Reference
Pro/Ser	311	0.0002[-0.0020,0.0024]	0.0005[-0.0054,0.0064]	0.0007[-0.0207,0.0222]	-0.0029[-0.0403,0.0344]
Ser/Ser	50	-0.0016[-0.0044,0.0012]	-0.0044[-0.0120,0.0033]	-0.0136[-0.0432,0.0160]	-0.0194[-0.0804,0.0416]
Pro/Ser + Ser/Ser	361	-0.0005[-0.0024,0.0014]	-0.0013[-0.0063,0.0038]	-0.004[-0.0229,0.0149]	-0.0075[-0.0416,0.0267]
rs10517 (3´ UTR (C/T))					
CC	584	Reference	Reference	Reference	Reference
CT	257	0.0009[-0.0014,0.0031]	0.0026[-0.0034,0.0086]	0.0095[-0.0130,0.0321]	0.0065[-0.0329,0.0459]
TT	24	0.0024[-0.0062,0.0111]	0.0060[-0.0155,0.0276]	0.0084[-0.0605,0.0772]	0.0084[-0.0999,0.1167]
CT + TT	281	0.001[-0.0012,0.0032]	0.0029[-0.0029,0.0087]	0.0097[-0.0120,0.0314]	0.0077[-0.0300,0.0453]

4.3.5 Impact of SNPs in Phase I Xenobiotic Genes on Air Pollution, Lung Function Interactions - Exploratory Analysis

Table 4.10 illustrates the genotype and allele distributions for the examined SNPs in AhR and CYP1A1 for the total population and stratified by ethnicity. The genotype distributions of the AhR SNPs were similar for all ethnic groups except for SNPs rs2066853, rs17722841 and rs2282885. The 'Arg" allele of rs2066853 (Arg/Lys) was the major allele in the Asian (67%) and White (79%), but not the Black (34%) population. Whereas for SNP rs17722841 (G/A), although 'G' was the major allele for all ethnicities, the genotype frequency varied between the White (67%) and other ethnic groups: Asian 95%, Black 95%, and 'Others' 84%. For SNP rs2282885 (T/C), 'T' was the major allele for all ethnic groups, however, with a different genotype distribution. Asian and Black populations had a similar genotype frequency of 'TT', 68% and 89% respectively, whilst White and 'Others' had a 39% and 44% genotype frequency respectively. Only one SNP of CYP1A1 displayed a similar minor allele frequency between the ethnicities, rs17861115 with the MAF varying between 4-12%. For SNP rs2606345 (G/T), although 'G' was the major allele for Asian and Black population, the genotype distribution was very different: 46% and 86% respectively. In contrast, 'TT' was the common genotype for White population, 64%. The allele 'G' was the major allele for all ethnicities for SNP rs2198843 (G/C) with the exception of the Black population where 'C' was the major allele. Whilst, Asian and 'Others' had a similar allele frequency of 'G', 63% and 67% respectively, the allele frequency was higher in White population (81%).

Tables I8 and **I9** in **Appendix I** show the crude and adjusted effects of the studied AhR and CYP1A1 genotypes on lung function. While unadjusted estimates suggested significant associations of children genotypes of AhR: rs2066853, rs17722841, rs2282885 and CYP1A1: rs2606345, rs2198843 with both FEV_1 and FVC, this effect could not be confirmed after adjustment for possible confounders, indicating that SNPs of AhR and CYP1A1 do not have a direct effect on lung function.

Table 4.10 Distribution of genotype and allele frequencies in AhR and CYP1A1

Genotype/Allele Frequencies	Ethnicity									
	Asian		Black		White		Others		Combined Ethnicity	
	n	%	n	%	n	%	n	%	n	%
AhR										
rs2074113 C/A										
CC	235	0.68	174	0.70	216	0.82	85	0.75	710	0.73
CA	102	0.29	67	0.27	45	0.17	25	0.22	239	0.25
AA	9	0.03	6	0.02	3	0.01	3	0.03	21	0.02
Allele C	572	0.83	415	0.84	477	0.90	195	0.86	1659	0.86
Allele A	120	0.17	79	0.16	51	0.10	31	0.14	281	0.14
rs2066853 Exon 10 Arg554Lys (G/A)										
Arg/Arg	232	0.67	83	0.34	211	0.79	63	0.55	589	0.61
Arg/Lys	105	0.30	113	0.46	50	0.19	45	0.39	313	0.32
Lys/Lys	9	0.03	51	0.21	5	0.02	6	0.05	71	0.07
Allele Arg	569	0.82	279	0.56	472	0.89	171	0.75	1491	0.77
Allele Lys	123	0.18	215	0.44	60	0.11	57	0.25	455	0.23
rs17722841 G/A										
GG	328	0.95	235	0.95	177	0.67	95	0.84	835	0.86
GA	19	0.05	12	0.05	83	0.31	17	0.15	131	0.13
AA	0	0.00	0	0.00	6	0.02	1	0.01	7	0.01
Allele G	675	0.97	482	0.98	437	0.82	207	0.92	1801	0.93
Allele A	19	0.03	12	0.02	95	0.18	19	0.08	145	0.07
rs17779352 T/C										
TT	278	0.80	230	0.93	222	0.83	104	0.92	834	0.86
TC	63	0.18	17	0.07	43	0.16	9	0.08	132	0.14
CC	6	0.02	0	0.00	1	0.00	0	0.00	7	0.01
Alelle T	619	0.89	477	0.97	487	0.92	217	0.96	1800	0.92
Allele C	75	0.11	17	0.03	45	0.08	9	0.04	146	0.08
rs2282885 T/C										
TT	233	0.68	219	0.89	103	0.39	50	0.44	605	0.62
TC	101	0.29	27	0.11	126	0.48	56	0.50	310	0.32
CC	11	0.03	1	0.00	36	0.14	7	0.06	55	0.06
Alelle T	567	0.82	465	0.94	332	0.63	156	0.69	1520	0.78
Allele C	123	0.18	29	0.06	198	0.37	70	0.31	420	0.22
CYP1A1										
rs2606345 G/T										
GG	161	0.46	213	0.86	34	0.13	33	0.29	441	0.45
GT	146	0.42	31	0.13	123	0.46	54	0.47	354	0.36
TT	40	0.12	3	0.01	109	0.41	27	0.24	179	0.18
Allele G	468	0.67	457	0.93	191	0.36	120	0.53	1236	0.63
Allele T	226	0.33	37	0.07	341	0.64	108	0.47	712	0.37
rs1799814 Exon 7 Thr461Asn										
Thr/Thr	330	0.97	240	0.97	234	0.88	101	0.90	905	0.94
Thr/Asn	8	0.02	7	0.03	28	0.11	10	0.09	53	0.05
Asn/Asn	1	0.00	0	0.00	4	0.02	1	0.01	6	0.01
Allele Thr	668	0.99	487	0.99	496	0.93	212	0.95	1863	0.97
Allele Asn	10	0.01	7	0.01	36	0.07	12	0.05	65	0.03
rs17861115 C/T										
CC	271	0.78	224	0.91	245	0.92	102	0.89	842	0.86
CT	68	0.20	22	0.09	20	0.08	10	0.09	120	0.12
TT	8	0.02	1	0.00	1	0.00	2	0.02	12	0.01
Allele C	610	0.88	470	0.95	510	0.96	214	0.94	1804	0.93
Allele T	84	0.12	24	0.05	22	0.04	14	0.06	144	0.07
rs2198843 G/C										
GG	142	0.41	45	0.18	177	0.67	55	0.48	419	0.43
GC	155	0.45	121	0.49	78	0.29	43	0.38	397	0.41
CC	50	0.14	80	0.33	11	0.04	16	0.14	157	0.16
Allele G	439	0.63	211	0.43	432	0.81	153	0.67	1235	0.63
Allele C	255	0.37	281	0.57	100	0.19	75	0.33	711	0.37

SNP which is not in HWE by ethnicity and total population.
Minor allele for this ethnicity is different with other ethnicities for the SNP.
Genotype distribution for this ethnicity are not similar with other group despite sharing same major/minor allele with the majority of the ethnici
Different getnoype distribution compared to other ethnicity.
Similar genotype/allele frequencies within the block for the particular SNP.
Similar genotype/allele frequencies within the block for the particular SNP.

Next, the effects of the selected SNPs in AhR and CYP1A1 on the association between the modeled air pollution exposures and lung function were examined. The interactions between Ahr genotypes and lung function for all annual pollutant exposures are shown in **Tables 4.11** (FEV_1) **- 4.12** (FVC), with the acute and sub-chronic exposure interactions presented in **Appendix I, Tables I10-I13**. I observed no evidence of an interaction of any of the Ahr SNPs on the relationship between FEV_1 (**Table 4.11**) or FVC (**Table 4.12)** with the annual modeled pollutant exposures, demonstrating that genetic variants in this gene do not influence the association between annual NOx exposures and reduced FVC. A nominally significant association between rs17722841 genotypes and FEV_1 in relation to sub-chronic NOx (7 day) exposures was also observed (**Table I12**), and with all pollutants when acute 24h exposures were considered (**Table I10**). However, none of these associations survived the multiple testing corrections.

The impacts of genetic variants in CYP1A1 on the relationship between long term exposure to air pollutants and FEV1 and FVC are presented in **Tables 4.13** and **4.14,** respectively; with additional information on the acute and sub-chronic exposures in **Appendix I**, **Tables I14-I17**. For CYP1A1, rs2198843 there was an apparent protective effect on lung function (FVC) in relation to both the sub-chronic and acute exposure attributions to both NOx and NO_2 that was robust to the multiple testing procedures, when the full multiethnic population was examined (**Figure 4.4**). Compared to children carrying wild type GG, children carrying the homozygous CC genotype had an increase in FVC of 5.2ml (2.6, 7.9; p=0.0013) per $\mu g/m^3$ and 14.0ml (7.1, 21.0; p=0.0008) per $\mu g/m^3$ in relation to average 7 day exposures to NOx and NO_2 respectively. Slightly smaller protective increments were observed in subjects with the homozygous genotype when the acute 24h exposures were considered: FVC 1.8ml (0.6, 3.0; p=0.0032) per $\mu g/m^3$ and 6.6ml (2.4, 10.7; p=0.0032) per $\mu g/m^3$ with respect to NOx and NO_2 exposure (**Figure 4.4**). A similar pattern was observed with the annual exposures, but this failed to attain statistical significance.

Table 4.11 Effects modification of AhR genotypes on the association between FEV_1 and annual mean pollutant exposures

Genotype	Subjects (n)	NOx 20m (n)	NO_2 20m β(95% CI)	PM_{10} 20m β (95% CI)	$PM_{2.5}$ 20m β (95% CI)
rs2074113 C/A					
CC	624	Reference	Reference	Reference	Reference
CA	218	-0.0004[-0.0021,0.0012]	-0.0015[-0.0060,0.0030]	-0.0049[-0.0231,0.0132]	-0.0105[-0.0444,0.0235]
AA	20	-31.0E-06[-0.0079,0.0079]	-0.0002[-0.0200,0.0196]	-0.0315[-0.0991,0.0360]	-0.0577[-0.1755,0.0601]
CA+AA	238	-0.0004[-0.0020,0.0012]	-0.0015[-0.0058,0.0029]	-0.0063[-0.0239,0.0113]	-0.0133[-0.0462,0.0195]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	516	Reference	Reference	Reference	Reference
Arg/Lys	281	-0.0001[-0.0019,0.0018]	-0.0005[-0.0054,0.0043]	-0.0004[-0.0187,0.0180]	-0.0175[-0.0508,0.0159]
Lys/Lys	68	0.0004[-0.0018,0.0026]	0.0008[-0.0052,0.0068]	0.0041[-0.0210,0.0291]	-0.0020[-0.0495,0.0455]
Arg/Lys+Lys/Lys	349	30.2E-06[-0.0016,0.0016]	-0.0003[-0.0046,0.0041]	0.0004[-0.0161,0.0169]	-0.0147[-0.0447,0.0153]
rs17722841 G/A					
GG	742	Reference	Reference	Reference	Reference
GA	116	-0.0005[-0.0038,0.0028]	-0.0010[-0.0095,0.0075]	-0.0055[-0.0344,0.0233]	-0.0113[-0.0577,0.0350]
AA	7	-0.0095[-0.0450,0.0260]	-0.0164[-0.0988,0.0660]	-0.0431[-0.3422,0.2560]	-0.0847[-0.3756,0.2061]
GA+AA	123	-0.0005[-0.0037,0.0027]	-0.0012[-0.0095,0.0071]	-0.006[-0.0344,0.0223]	-0.0132[-0.0586,0.0322]
rs17779352 T/C					
TT	746	Reference	Reference	Reference	Reference
TC	115	0.0009[-0.0015,0.0033]	0.0025[-0.0039,0.0090]	0.0054[-0.0180,0.0288]	0.0298[-0.0145,0.0741]
CC	4	-0.0192[-0.0488,0.0104]	-0.0456[-0.1167,0.0255]	-0.1665[-0.4473,0.1143]	-0.2639[-0.6246,0.0969]
TC+CC	119	0.0007[-0.0016,0.0031]	0.0021[-0.0042,0.0085]	0.0042[-0.0189,0.0274]	0.0263[-0.0173,0.0700]
rs2282885 T/C					
TT	547	Reference	Reference	Reference	Reference
TC	264	-0.0002[-0.0020,0.0016]	-0.0005[-0.0055,0.0045]	-0.0006[-0.0194,0.0183]	0.0004[-0.0326,0.0334]
CC	51	-0.0007[-0.0080,0.0065]	-0.0019[-0.0204,0.0166]	-0.0336[-0.0978,0.0305]	-0.0284[-0.1074,0.0506]
TC+CC	315	-0.0001[-0.0019,0.0017]	-0.0003[-0.0051,0.0045]	-0.0018[-0.0200,0.0164]	-0.0023[-0.0337,0.0290]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.

Table 4.12 Effects modification of AhR Genotypes on association between FVC and annual mean pollutant exposures

Genotype	Subjects (n)	NOx 20m (n)	NO_2 20m β (95% CI)	PM_{10} 20m β (95% CI)	$PM_{2.5}$ 20m β (95% CI)
rs2074113 C/A					
CC	604	Reference	Reference	Reference	Reference
CA	210	-0.0002[-0.0021,0.0017]	-0.0011[-0.0062,0.0040]	-0.0034[-0.0242,0.0174]	-0.0099[-0.0490,0.0293]
AA	20	0.0045[-0.0045,0.0134]	0.0119[-0.0104,0.0343]	0.0161[-0.0604,0.0927]	-0.0039[-0.1374,0.1297]
CA+AA	230	-0.0001[-0.0019,0.0018]	-0.0006[-0.0056,0.0044]	-0.0025[-0.0227,0.0178]	-0.0096[-0.0474,0.0283]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	498	Reference	Reference	Reference	Reference
Arg/Lys	270	0.0002[-0.0019,0.0023]	-42.5E-06[-0.0056,0.0055]	0.0023[-0.0188,0.0233]	-0.0050[-0.0435,0.0336]
Lys/Lys	68	0.0008[-0.0017,0.0033]	0.0023[-0.0045,0.0090]	0.0129[-0.0156,0.0413]	0.0108[-0.0433,0.0649]
Arg/Lys+Lys/Lys	338	0.0004[-0.0015,0.0022]	0.0006[-0.0043,0.0056]	0.005[-0.0139,0.0239]	-0.0016[-0.0361,0.0329]
rs17722841 G/A					
GG	717	Reference	Reference	Reference	Reference
GA	113	-0.0011[-0.0049,0.0028]	-0.0027[-0.0126,0.0072]	-0.0098[-0.0440,0.0244]	-0.0134[-0.0681,0.0412]
AA	7	-0.0061[-0.0464,0.0342]	-0.0092[-0.1026,0.0842]	-0.0107[-0.3500,0.3287]	-0.0285[-0.3586,0.3017]
GA+AA	120	-0.001[-0.0048,0.0027]	-0.0026[-0.0123,0.0071]	-0.0095[-0.0431,0.0240]	-0.0136[-0.0670,0.0399]
rs17779352 T/C					
TT	722	Reference	Reference	Reference	Reference
TC	110	30.4E-06[-0.0028,0.0028]	0.0001[-0.0075,0.0077]	-0.0017[-0.0288,0.0255]	0.0031[-0.0487,0.0550]
CC	4	-0.0148[-0.0484,0.0187]	-0.0351[-0.1157,0.0455]	-0.1313[-0.4495,0.1870]	-0.1977[-0.6071,0.2118]
TC+CC	114	-46.7E-06[-0.0028,0.0027]	-0.0001[-0.0076,0.0073]	-0.0025[-0.0293,0.0244]	0.0002[-0.0509,0.0512]
rs2282885 T/C					
TT	528	Reference	Reference	Reference	Reference
TC	257	-0.0002[-0.0023,0.0019]	-0.0009[-0.0066,0.0047]	-0.0041[-0.0257,0.0176]	-0.0030[-0.0410,0.0350]
CC	48	0.0010[-0.0075,0.0094]	0.0032[-0.0183,0.0247]	-0.0133[-0.0921,0.0656]	-0.0161[-0.1072,0.0751]
TC+CC	305	-2.20E-06[-0.0020,0.0020]	-0.0002[-0.0057,0.0052]	-0.003[-0.0240,0.0179]	-0.0032[-0.0393,0.0329]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.

Table 4.13 Effect modification of CYP1A1 genotypes on the association between FEV_1 and annual mean pollutant exposures

Genotype	Subjects (n)	NOx 20m (n)	NO_2 20m β(95% CI)	PM_{10} 20m β (95% CI)	$PM_{2.5}$ 20m β (95% CI)
rs2606345 G/T					
GG	391	Reference	Reference	Reference	Reference
GT	319	-0.0024**[-0.0041,-0.0006]	-0.0064**[-0.0112,-0.0016]	-0.0217*[-0.0401,-0.0034]	-0.0213[-0.0542,0.0117]
TT	156	-0.0002[-0.0026,0.0021]	-0.0011[-0.0073,0.0051]	-0.0045[-0.0273,0.0182]	0.0011[-0.0399,0.0420]
GT+TT	475	-0.0017*[-0.0034,-0.0001]	-0.0048*[-0.0093,-0.0004]	-0.0159#[-0.0326,0.0007]	-0.0136[-0.0432,0.0160]
rs17861115 C/T					
CC	749	Reference	Reference	Reference	Reference
CT	106	-0.0005[-0.0042,0.0033]	-0.0007[-0.0102,0.0087]	0.0014[-0.0334,0.0361]	0.0094[-0.0452,0.0640]
TT	11	0.0047[-0.0044,0.0138]	0.0121[-0.0108,0.0349]	0.0538[-0.0208,0.1284]	0.1218[-0.0281,0.2717]
CT+TT	117	0.0003[-0.0032,0.0038]	0.0012[-0.0075,0.0099]	0.0107[-0.0209,0.0423]	0.0216[-0.0296,0.0729]
rs2198843 G/C					
GG	373	Reference	Reference	Reference	Reference
GC	349	0.0001[-0.0019,0.0020]	0.0001[-0.0048,0.0051]	-0.0002[-0.0183,0.0180]	-0.0085[-0.0409,0.0240]
CC	144	0.0033*[0.0002,0.0065]	0.0087*[0.0007,0.0167]	0.0245#[-0.0046,0.0535]	0.011[-0.0369,0.0589]
GC+CC	493	0.0008[-0.0009,0.0026]	0.0022[-0.0024,0.0068]	0.0056[-0.0112,0.0225]	-0.0032[-0.0335,0.0270]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01.

Table 4.14 Effect modification of CYP1A1 genotypes on association between FVC and annual mean pollutant exposures

Genotype	Subjects (n)	NOx 20m (n)	NO_2 20m β (95% CI)	PM_{10} 20m β (95% CI)	$PM_{2.5}$ 20m β (95% CI)
rs2606345 G/T					
GG	380	Reference	Reference	Reference	Reference
GT	310	-0.0026*[-0.0046,-0.0005]	-0.0071*[-0.0126,-0.0016]	-0.0253*[-0.0463,-0.0043]	-0.0306[-0.0684,0.0072]
TT	147	-0.0003[-0.0030,0.0024]	-0.0015[-0.0086,0.0056]	-0.0047[-0.0309,0.0214]	0.0051[-0.0421,0.0522]
GT+TT	457	-0.0019*[-0.0038,-48.0E-6]	-0.0055*[-0.0106,-0.0004]	-0.0186#[-0.0377,0.0005]	-0.0188[-0.0529,0.0152]
rs17861115 C/T					
CC	723	Reference	Reference	Reference	Reference
CT	103	-0.0029[-0.0075,0.0018]	-0.0066[-0.0183,0.0051]	-0.0210[-0.0627,0.0207]	-0.0294[-0.0945,0.0356]
TT	11	0.0068[-0.0035,0.0171]	0.0177[-0.0081,0.0436]	0.0742#[-0.0102,0.1586]	0.1596#[-0.0101,0.3293]
CT+TT	114	-0.0012[-0.0055,0.0030]	-0.0025[-0.0132,0.0081]	-0.0028[-0.0403,0.0348]	-0.0063[-0.0671,0.0544]
rs2198843 G/C					
GG	354	Reference	Reference	Reference	Reference
GC	340	-0.0006[-0.0027,0.0016]	-0.0012[-0.0069,0.0046]	-0.0037[-0.0245,0.0170]	-0.0135[-0.0508,0.0237]
CC	143	0.0034[-0.0002,0.0070]	0.0094*[0.0002,0.0187]	0.0291#[-0.0050,0.0632]	0.0126[-0.0433,0.0685]
GC+CC	483	0.0004[-0.0016,0.0024]	0.0014[-0.0039,0.0066]	0.0037[-0.0157,0.0230]	-0.0068[-0.0416,0.0281]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.
#p<0.1, *p<0.05.

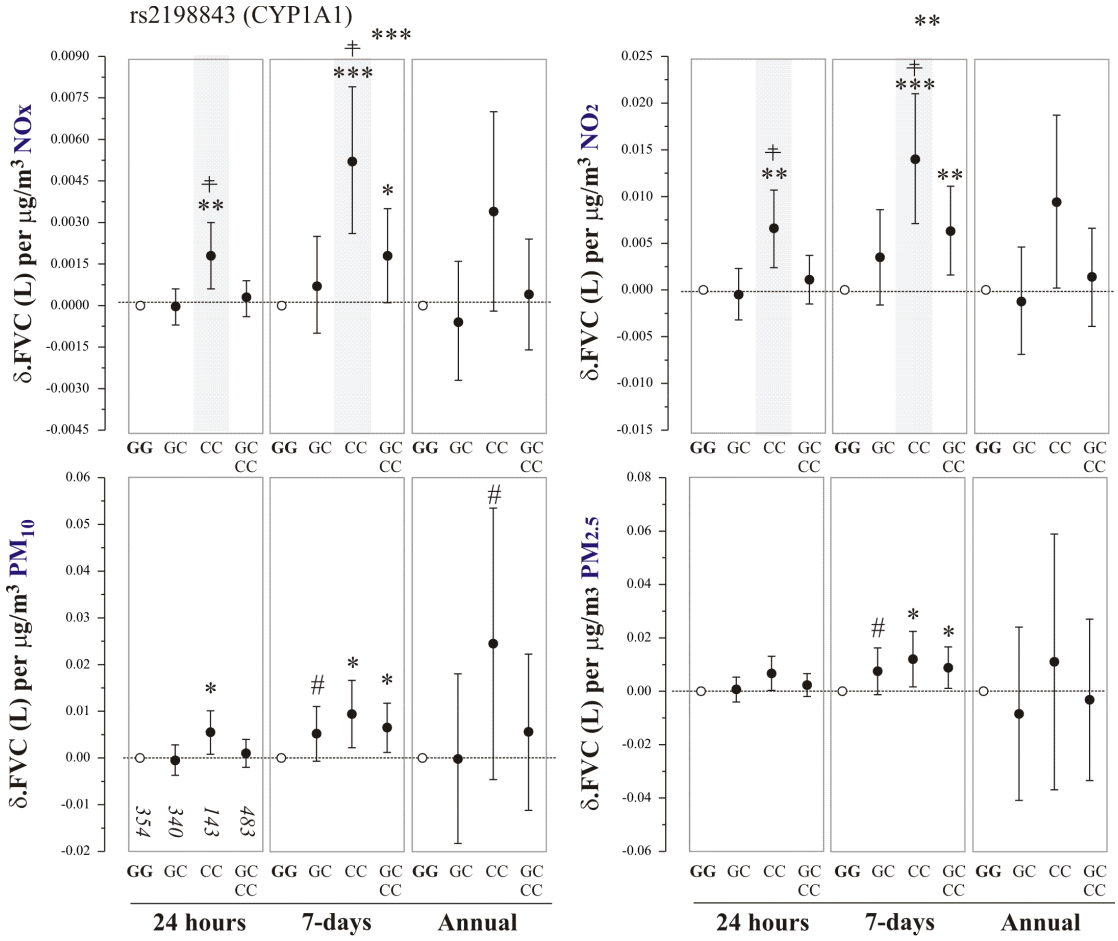


Figure 4.4 Influence of rs2198843 (CYP1A1) on the association between post-bronchodilator FVC and 24 hour, 7-day or annual pollutant attributions to NOx, NO_2, PM_{10} and $PM_{2.5}$. Data are expressed relative to the major allele across the whole multi-ethnic population- GG (**Table 4.10**). Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year, including effect for school location. Nominal significant differences relative to the reference genotype are illustrated as follows: #p<0.1, *p<0.05 and **p<0.01. Where these differences remained significant (P<0.05) after adjusted for multiple comparisons this is illustrated by '†'.

The analyses were repeated separately for the three major ethnic groups within the present study. This was necessary for rs2198843, as the major allele differed between the Asian and White (G, major allele fregency 63 and 81% respectively), versus the black populations (C, major allele fregency 57%). The results of these separate analyses are summarised in **Table 4.15**. Overall these data supported the view that possession of the 'C' allele is protective against

pollutant associated reductions in FVC, but this effect was only strongly evident in the Asian and black children and was predmoniately related to the acute (24h) to sub-chronic (7-day) exposures. For the Asian children the protective effect was seen in individuals homozygous for the C allele for each pollutant examined, and in the combined models (GC+CC) for 7 day average PM_{10} and $PM_{2.5}$ exposures. In contrast, in the black children possession of the G allele was associated with an increased negative association between acute (for all pollutants) and sub-chronic (NOx and NO_2) pollutant exposures and FVC. The magnitude of the effect increased from heterozygotes to homozygotes for the acute 24h exposures, but was only evident in the heterozygotes and in the combined model (CG+GG) using the 7 day exposure estimates, likely reflecting the small number of homozygous subjects (n=39).

In contrast, an increased decrement in FVC in relation to sub-chronic and acute exposures to NOx and NO_2 was observed with rs2606345, in both the combined and heterozygous genotype when the whole population was considered together (**Figure 4.5**). Children carrying heterozygous genotype (GT) demonstrated a significant decrement in FVC related to their sub-chronic exposure to NOx and NO_2, relative to the dominant GG genotype (63%). When these data were analysed separately by ethnicity, with reference to each populations major allele: G for the Asian (67%) and black (93%) populations and T for the white children (64%), the sensitizing affect of possession of the T allele was still apparent, but as for for rs2198843, this was limited to the Asian and black children, **Table 4.16**. For the Asian children possession of the T allele was associated with a significant negative impact of the underlying association between sub-chronic exposures to NOx, NO_2 and PM_{10} in the heterozygotes and in the combined model, but not in the homozygotes, again likely reflecting the small number of subjects (n=29). For the black children this addative negative impact was only apparent in relation to 7-day exposures to PM_{10}.

Table 4.15 Influence of rs2198843 (CYP1A1) on the association between post-bronchodilator FVC and 24 hour, 7-day or annual pollutant attributions to NOx, NO_2, PM_{10} and $PM_{2.5}$ for each of the major ethnicities in the current population. Data are expressed relative to the major allele for each ethnic population

Genotype	Subjects	NOx 24 hours		NO_2 24 hours		PM_{10} 24 hours		$PM_{2.5}$ 24 hours	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/C)									
GG	132	*Reference*		*Reference*		*Reference*		*Reference*	
GC	45	0.0000	[-0.0016,0.0016]	-0.0006	[-0.0056,0.0044]	-0.0023	[-0.0081,0.0034]	-0.0014	[-0.0106,0.0079]
CC	290	0.0006	[-0.0015,0.0027]	0.0044	[-0.0029,0.0117]	0.0014	[-0.0063,0.0091]	0.0011	[-0.0101,0.0123]
GC+CC	177	0.0001	[-0.0014,0.0017]	0.0006	[-0.0042,0.0053]	-0.0014	[-0.0067,0.0040]	-0.0006	[-0.0093,0.0081]
Black (C/G)									
CC	74	*Reference*		*Reference*		*Reference*		*Reference*	
CG	104	**-0.0022***	**[-0.0041,-0.0004]**	**-0.0065***	**[-0.0130,-0.0000]**	**-0.0094***	**[-0.0171,-0.0016]**	-0.0091	[-0.0201,0.0018]
GG	39	**-0.0035****	**[-0.0060,-0.0010]**	**-0.0096***	**[-0.0183,-0.0010]**	**-0.0156****	**[-0.0256,-0.0055]**	**-0.0171***	**[-0.0307,-0.0035]**
CG+GG	143	**-0.0025****	**[-0.0043,-0.0006]**	**-0.0073***	**[-0.0136,-0.0010]**	**-0.0106****	**[-0.0181,-0.0030]**	**-0.0109***	**[-0.0213,-0.0004]**
White (G/C)									
GG	159	***Reference***		***Reference***		***Reference***		***Reference***	
GC	69	-0.0001	[-0.0010,0.0008]	-0.0001	[-0.0044,0.0042]	-0.0007	[-0.0057,0.0042]	-0.0005	[-0.0075,0.0064]
CC	9	0.002	[-0.0013,0.0054]	0.008	[-0.0047,0.0208]	0.0087	[-0.0036,0.0209]	0.0113	[-0.0063,0.0290]
GC+CC	78	0.000	[-0.0009,0.0009]	0.0003	[-0.0039,0.0046]	0.0001	[-0.0047,0.0048]	0.0006	[-0.0062,0.0073]

Genotype	Subjects	NOx 7 days		NO_2 7 days		PM_{10} 7 days		$PM_{2.5}$ 7 days	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/C)									
GG	132	*Reference*		*Reference*		*Reference*		*Reference*	
GC	45	0.0012	[-0.0016,0.0040]	0.0035	[-0.0047,0.0118]	0.0055	[-0.0016,0.0125]	0.0094	[-0.0016,0.0204]
CC	290	**0.0049***	**[0.0005,0.0093]**	**0.0164***	**[0.0039,0.0289]**	**0.0134***	**[0.0025,0.0243]**	**0.0166***	**[0.0011,0.0320]**
GC+CC	177	0.0019	[-0.0007,0.0046]	0.0064	[-0.0014,0.0142]	**0.0072***	**[0.0006,0.0138]**	**0.0112***	**[0.0010,0.0213]**
Black (C/G)									
CC	74	*Reference*		*Reference*		*Reference*		*Reference*	
CG	104	**-0.0054***	**[-0.0096,-0.0013]**	**-0.0109***	**[-0.0214,-0.0003]**	-0.0052	[-0.0191,0.0087]	-0.0083	[-0.0289,0.0124]
GG	39	-0.005	[-0.0110,0.0010]	-0.0103	[-0.0258,0.0052]	-0.0139	[-0.0309,0.0031]	-0.0179	[-0.0418,0.0061]
CG+GG	143	**-0.0054****	**[-0.0095,-0.0014]**	**-0.0111***	**[-0.0214,-0.0009]**	-0.0074	[-0.0205,0.0056]	-0.0104	[-0.0291,0.0083]
White (G/C)									
GG	159	***Reference***		***Reference***		***Reference***		***Reference***	
GC	69	0.002	[-0.0013,0.0053]	0.0074	[-0.0020,0.0169]	-0.0004	[-0.0139,0.0131]	-0.0029	[-0.0223,0.0165]
CC	9	0.0021	[-0.0094,0.0137]	0.0048	[-0.0253,0.0349]	0.0232	[-0.0164,0.0628]	0.0297	[-0.0170,0.0764]
GC+CC	78	0.002	[-0.0012,0.0052]	0.0073	[-0.0019,0.0166]	0.0007	[-0.0124,0.0139]	0.0009	[-0.0178,0.0195]

Genotype	Subjects	NOx annual		NO_2 24 annual		PM_{10} annual		$PM_{2.5}$ annual	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/C)									
GG	132	*Reference*		*Reference*		*Reference*		*Reference*	
GC	45	-0.0007	[-0.0041,0.0028]	-0.0015	[-0.0104,0.0074]	0.0009	[-0.0328,0.0346]	-0.0082	[-0.0717,0.0552]
CC	290	0.0055	[-0.0004,0.0114]	0.015	[-0.0001,0.0301]	0.039	[-0.0200,0.0980]	0.024	[-0.0703,0.1183]
GC+CC	177	0.0008	[-0.0024,0.0039]	0.0022	[-0.0059,0.0104]	0.0087	[-0.0227,0.0401]	-0.0004	[-0.0598,0.0590]
Black (C/G)									
CC	74	*Reference*		*Reference*		*Reference*		*Reference*	
CG	104	-0.0023	[-0.0084,0.0038]	-0.0061	[-0.0218,0.0096]	-0.0237	[-0.0787,0.0313]	-0.0105	[-0.0987,0.0778]
GG	39	-0.002	[-0.0089,0.0049]	-0.0046	[-0.0224,0.0133]	-0.0249	[-0.0878,0.0380]	0.0012	[-0.1128,0.1153]
CG+GG	143	-0.0022	[-0.0080,0.0036]	-0.0056	[-0.0205,0.0094]	-0.0232	[-0.0761,0.0296]	-0.0052	[-0.0910,0.0805]
White (G/C)									
GG	159	*Reference*		*Reference*		*Reference*		*Reference*	
GC	69	0.0004	[-0.0035,0.0042]	0.0022	[-0.0077,0.0121]	0.0059	[-0.0298,0.0417]	-0.0164	[-0.0794,0.0466]
CC	9	0.0012	[-0.0150,0.0174]	0.0051	[-0.0354,0.0457]	0.0319	[-0.1073,0.1712]	0.1078	[-0.0869,0.3025]
GC+CC	78	0.0003	[-0.0035,0.0041]	0.002	[-0.0078,0.0118]	0.0058	[-0.0295,0.0410]	-0.0075	[-0.0693,0.0542]

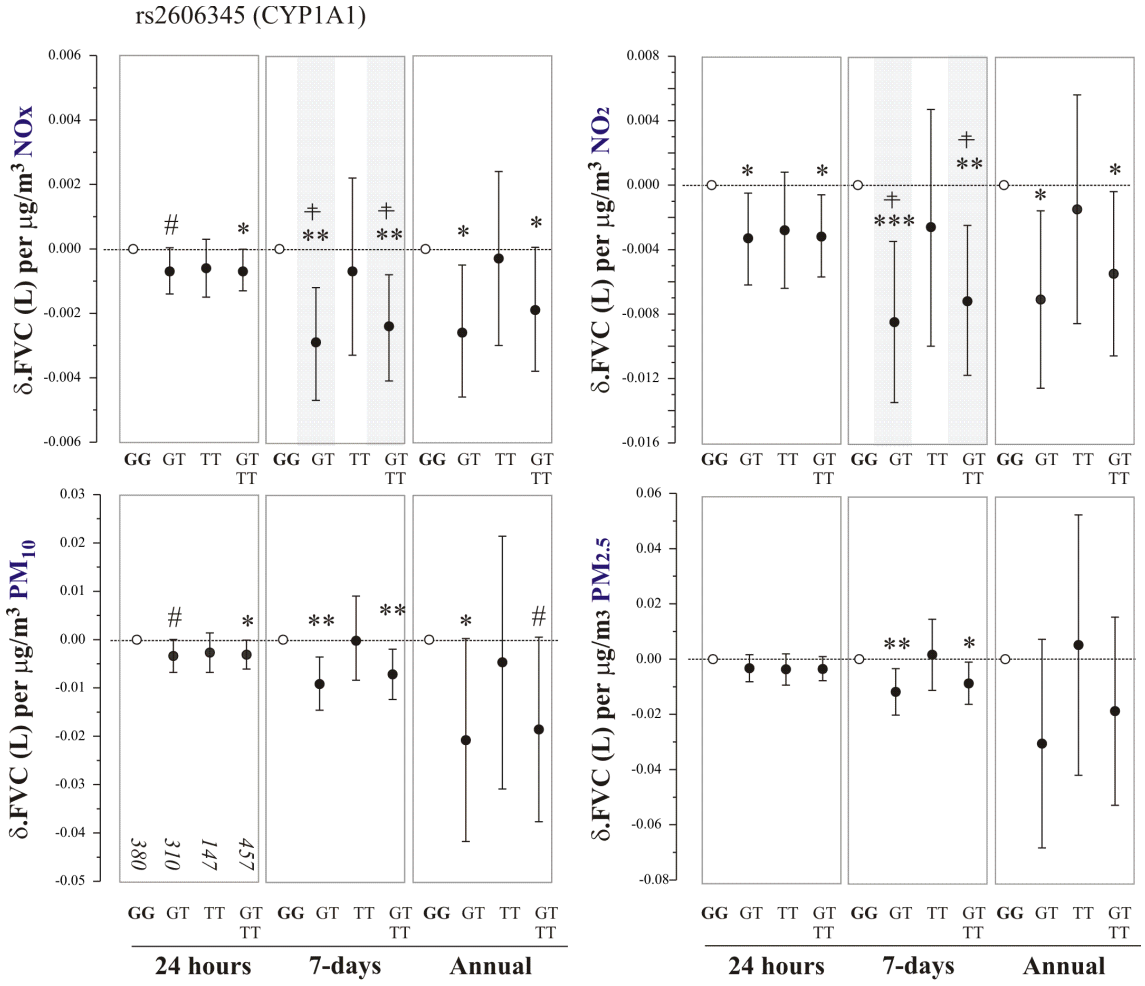


Figure 4.5 Influence of rs2606345 (CYP1A1) on the association between post-bronchodilator FVC and 24 hour, 7-day or annual pollutant attributions to NOx, NO_2, PM_{10} and $PM_{2.5}$. Data are expressed relative to the major allele across the whole multi-ethnic population- GG (**Table 4.10**). Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year, including effect for school location. Nominal significant differences relative to the reference genotype are illustrated as follows: #p<0.1, *p<0.05 and **p<0.01. Where these differences remained significant (P<0.05) after adjusted for multiple comparisons this is illustrated by '†'.

Table 4.16 Influence of rs2606345 (CYP1A1) on the association between post-bronchodilator FVC and 24 hour, 7-day or annual pollutant attributions to NOx, NO_2, PM_{10} and $PM_{2.5}$ for each of the major ethnicities in the current population. Data are expressed relative to the major allele for each ethnic population

Genotype	Subjects	NOx 24 hours		NO_2 24 hours		PM_{10} 24 hours		$PM_{2.5}$ 24 hours	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/T)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GT	123	-0.0009	[-0.0023,0.0005]	**-0.0048***	**[-0.0096,-0.0000]**	-0.0029	[-0.0086,0.0027]	-0.0019	[-0.0105,0.0066]
TT	29	-0.0003	[-0.0029,0.0023]	-0.0016	[-0.0095,0.0063]	-0.0003	[-0.0093,0.0087]	0.0001	[-0.0154,0.0156]
GT+TT	152	-0.0008	[-0.0021,0.0005]	-0.0042	[-0.0088,0.0003]	-0.0022	[-0.0075,0.0030]	-0.0013	[-0.0094,0.0067]
Black (G/T)									
GG	187	*Reference*		*Reference*		*Reference*		*Reference*	
GT	29	-0.0024	[-0.0052,0.0005]	-0.0064	[-0.0165,0.0037]	-0.0101	[-0.0206,0.0003]	-0.0109	[-0.0267,0.0049]
TT	1	.	.	.	.	.	.	.	.
GT+TT	30	-0.0024	[-0.0052,0.0005]	-0.0064	[-0.0164,0.0035]	-0.01	[-0.0202,0.0003]	-0.0107	[-0.0260,0.0047]
White (T/G)									
TT	98	*Reference*		*Reference*		*Reference*		*Reference*	
TG	113	-0.0008	[-0.0027,0.0011]	-0.0025	[-0.0095,0.0045]	-0.003	[-0.0110,0.0050]	-0.005	[-0.0171,0.0071]
GG	26	-0.0013	[-0.0032,0.0007]	-0.0041	[-0.0113,0.0032]	-0.0048	[-0.0129,0.0032]	-0.0084	[-0.0205,0.0037]
TG+GG	211	-0.001	[-0.0028,0.0009]	-0.0032	[-0.0099,0.0036]	-0.0039	[-0.0115,0.0038]	-0.0067	[-0.0184,0.0049]

Genotype	Subjects	NOx 7 days		NO_2 7 days		PM_{10} 7 days		$PM_{2.5}$ 7 days	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/T)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GT	123	**-0.0039****	**[-0.0066,-0.0012]**	**-0.0145*****	**[-0.0225,-0.0065]**	**-0.0084***	**[-0.0153,-0.0016]**	-0.0102	[-0.0208,0.0003]
TT	29	-0.0003	[-0.0049,0.0044]	0.0002	[-0.0132,0.0136]	-0.0017	[-0.0189,0.0156]	-0.0061	[-0.0327,0.0204]
GT+TT	152	**-0.0034***	**[-0.0059,-0.0008]**	**-0.0122****	**[-0.0199,-0.0044]**	**-0.0079***	**[-0.0147,-0.0012]**	-0.01	[-0.0203,0.0004]
Black (G/T)									
GG	187	*Reference*		*Reference*		*Reference*		*Reference*	
GT	29	-0.001	[-0.0069,0.0049]	-0.0026	[-0.0173,0.0121]	**-0.0221***	**[-0.0421,-0.0021]**	-0.0241	[-0.0511,0.0028]
TT	1	.		.		.		.	
GT+TT	30	-0.001	[-0.0069,0.0049]	-0.0026	[-0.0173,0.0120]	**-0.0222***	**[-0.0421,-0.0022]**	-0.0241	[-0.0508,0.0026]
White (T/G)									
TT	98	*Reference*		*Reference*		*Reference*		*Reference*	
TG	113	-0.0012	[-0.0067,0.0043]	-0.0023	[-0.0171,0.0125]	-0.0062	[-0.0277,0.0153]	-0.0169	[-0.0455,0.0117]
GG	26	-0.0002	[-0.0061,0.0057]	-0.0013	[-0.0170,0.0144]	0.0051	[-0.0169,0.0272]	-0.0026	[-0.0314,0.0262]
TG+GG	211	-0.0009	[-0.0063,0.0045]	-0.0019	[-0.0164,0.0125]	-0.0012	[-0.0220,0.0196]	-0.01	[-0.0373,0.0174]

Genotype	Subjects	NOx annual		NO_2 24 annual		PM_{10} annual		$PM_{2.5}$ annual	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/T)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GT	123	-0.0028	[-0.0056,0.0000]	**-0.0079***	**[-0.0157,-0.0001]**	-0.0237	[-0.0571,0.0098]	0.0019	[-0.0594,0.0632]
TT	29	0.0001	[-0.0037,0.0039]	-0.0006	[-0.0112,0.0099]	-0.004	[-0.0498,0.0418]	0.0317	[-0.0617,0.1251]
GT+TT	152	-0.0021	[-0.0048,0.0006]	-0.0062	[-0.0137,0.0013]	-0.0193	[-0.0517,0.0130]	0.0088	[-0.0498,0.0674]
Black (G/T)									
GG	187	*Reference*		*Reference*		*Reference*		*Reference*	
GT	29	-0.0077	[-0.0175,0.0021]	-0.0195	[-0.0442,0.0051]	-0.0666	[-0.1487,0.0154]	-0.1174	[-0.2390,0.0042]
TT	1	.		.		.		.	
GT+TT	30	-0.0077	[-0.0176,0.0021]	-0.0196	[-0.0442,0.0051]	-0.0667	[-0.1488,0.0154]	-0.1177	[-0.2386,0.0033]
White (T/G)									
TT	98	*Reference*		*Reference*		*Reference*		*Reference*	
TG	113	0.0015	[-0.0047,0.0076]	0.0039	[-0.0125,0.0202]	-0.0003	[-0.0498,0.0493]	-0.0157	[-0.1096,0.0783]
GG	26	0.0032	[-0.0034,0.0098]	0.0078	[-0.0097,0.0253]	0.019	[-0.0328,0.0709]	0.0202	[-0.0763,0.1167]
TG+GG	211	0.0019	[-0.0042,0.0080]	0.005	[-0.0112,0.0212]	0.0075	[-0.0401,0.0552]	-0.0002	[-0.0908,0.0904]

4.4 DISCUSSION

Roughly half of the world's population - 3.4 billion people - live in urban areas (United Nations 2009). By 2030, 60% of the global population are projected to live in cities (van Ginkel et al. 2007), and with increasing population growth and continuing urbanization air quality has emerged as an important determinant both of the global burden of disease (Lim et al. 2012) and of health within cities (Lepeule et al. 2012; Gauderman et al. 2007; Grigg 2012). Larger urban populations are associated with increased vehicle ownership and consequently greater emissions of traffic-derived pollutants and adverse air quality. Exposure to traffic emissions adversely affects both respiratory and cardiovascular health (HEI 2010). In addition, studies following children's lung function over time have demonstrated that traffic fumes damage the development of children's lungs (Gauderman et al. 2004, 2007). Children living close to main roads also report more respiratory symptoms such as a worsening of asthma symptoms and cough (Janssen et al. 2003; Hoek et al. 2012).

Interventions that aim to improve health by controlling or reducing traffic emissions are therefore increasingly important elements of city planning to maintain good public health in metropolitan areas. The introduction of low emission zones within cities, restricting the entry of the most polluting vehicles, or promoting the uptake of improved emission controls, has therefore been employed in an attempt to improve air quality (Kelly et al. 2011, 2012). In February 2008, London established such a Low Emission Zone, encompassing most of greater London, an area of 2,644km^2, representing the world's largest such initiative, impacting upon a resident population of 8,174,100 (2011 UK census), plus 790,000 (2009 TFL traffic survey) daily commuters from outside the city. The London scheme was designed to have progressive tightening of the emission standards applied to increased proportions of the vehicle fleet over time. The initial two phases in 2008 applied Euro III PM_{10} emission standards to heavy goods vehicles (over 12 tonnes – phase 1) and subsequently to all vehicles (3.5-12 tonnes) and buses/ coaches (phase 2). Feasibility studies predicted that these changes would deliver

a reduction in NOx and PM_{10} emission of 1288 and 81 tonnes respectively. Whilst the overall impact of these changes on air quality were predicted to be small throughout the London, significant benefits were predicted in high traffic areas (Kelly et al., 2011). As was stated previously, the present study did not begin until November 2008, after the introduction of the first two phases of the LEZ. Whilst these two phases were predicted to result in reductions in PM_{10}, the actual reduction was always going to be small and only measureable at defined hot spots. The present study is still ongoing and has just completed its fifth year of sampling, Nov 2012 – March 2013. In the interim period the phases III and IV of the LEZ were introduced in Jan 2012, with EURO IV emission standards applied to a broader range of vehicle types (all goods vehicles over 3.5 tonnes, including buses/coaches). This has been predicted to achieve reductions of 3,139 tonnes of NOx and 204 tonnes of PM_{10} respectively. Thus, the current data effectively constitutes a baseline to further interrogate the effectiveness of this traffic management scheme. In the present study, I therefore examined the relationship between lung function and acute to chronic pollutant exposures in children living in two inner-city boroughs within the LEZ; Tower Hamlets and Hackney. These areas are highly polluted, as reviewed in **Chapter 3**.

This study focused on children as a vulnerable group, with high pollutant burdens for both biologic and social reasons. Children have a larger lung-to-body volume ratio, their airway epithelium is more permeable to air pollutants, and their lung defense mechanisms against particulate matter (PM) and gaseous pollutants are not considered to be fully matured (Schwartz 2004). It has been estimated, based on these facts that a child breathing polluted urban air will receive up to a 2- to 4-fold higher dose than an adult (Ginsberg 2005). There is also a growing evidence base suggesting that the mother's exposure to air pollution during intrauterine development is associated with increased rates of infant mortality (Lipfert et al. 2000), prematurity (Sagiv et al. 2005), congenital defects (Gilboa et al. 2005), and lower birth weight (Lee et al. 2003; Dugandzic et al. 2006); as well as increased respiratory symptoms (Jedrychowski et al. 2005), decreased lung function (Turnovskaa and Marinov 2009) and poor cognitive development (Edwards et al. 2010) in early life.

Consistent with this perceived sensitivity small deficits in lung function and growth have been reported in children living near busy roads with high long-term exposures to traffic-related air pollutants (Jedrychowski et al. 1999; Peters et al. 1999; Horak et al. 2002; Gauderman et al. 2004, 2007; Rojas-Martinez et al. 2007; Nordling et al. 2008; Oftedal et al. 2008; He et al. 2010; Roy et al. 2012; Schultz et al. 2012; Eenhuisen et al. 2013). However, it is important to note that the literature for a negative association between air pollutant exposures in children and lung function is not wholly coherent and numerous studies have failed to report significant associations (Dockery et al. 1989; Hirsch et al. 1999; Nicolai et al. 2003), or have reported effects only in sub-groups (Peters et al. 1999). This literature has been reviewed extensively in recent times (Gotschi et al. 2008; HEI 2010), but due to variations in study design, exposure attribution (modelling versus measurement), the lung function variables used and the populations examined no simple clear message has emerged. A recent meta-analyis by Hoek et al. (2012) examined the relationship between PM_{10} exposure and lung function (FVC, FEV_1, PEF, FEF_{25-75}) in 22,809 children, based on historic cohorts in Linz, Austria (6-8 year old children, study period, Jan 1996 – Dec 1998 (Moshammer et al. 2006)), various study centres in central and Eastern Europe (9-12 year olds, Spring 1996 (Pattenden et al. 2000)); three towns in East Germany (6-12 years, September 1992 – July 1993 – (Heinrich et al. 1999)); 24 schools within the Netherlands (7-12 years, study period, April 1997 – July 1998 (Janssen et al. 2003)) and 24 cities in the USA and Canada (8-12 years, 1988-1991 (Dockery et al. 1996; Raizenne et al. 1996). In this analysis, they found no evidence of a relationship with PM_{10} and lung function, consistent with previous studies (Dockery et al. 1989; Hirsch et al. 1999; Nicolai et al. 2003).

An attempt to address discrepancies between study designs was performed as part of the recently completed European Study of Cohorts for Air Pollution Effects (ESCAPE) project. This study pooled equivalent lung function data collected from children (6-8 years) across five European cohorts, between 2002-2007 and related these to annual pollutant exposures based on residential address using Land Use Regression models. In addition, the influence of short term exposures on the lung function

measurements were assessed based on measurements made at regional and urban background monitoring sites (Gehring et al. 2013). This study, with a combined population of 5,921 children demonstrated significant reductions in FEV_1, associated with annual exposures to NO_2, NOx, $PM_{2.5}$ and $PM_{2.5}$ absorbance.

The present study therefore appears to confirm this observation and the bulk of the existing literature demonstrating negative associations between lung function and exposure to urban air, particularly to NOx and NO_2, good tracers of diesel emissions within London. The fact that this effect was only apparent in the post-bronchodilator FVC measurements and not FEV_1 and only with the annual and not the shorter term exposure attributions strongly imply that this result reflects reduced lung growth. The negative impact of primary traffic derived pollutants on children's lung function has recently been further reinforced in a US study, where traffic exposures (again estimated using Land Use Regression models) were associated with reduced FEV_1 and FVC in a cohort of 1,811 5-7 year old children (Urman et al. 2013). It is notable that compared with these recent studies the annual exposure to pollutants in our London cohort, particularly for NOx and NO_2, are very high; greater than the NO_2 concentrations in the high pollution communities (Anaheim, Glendora and San Dimas) in the study by Urman et al. (2013), and almost two-fold higher than the concentrations in the highest exposure cohort in the ESCAPE project (GIMI/LISA North): 43.52 ± 5.45 versus 23.4 ± 2.8 $\mu g/m^3$ (Gehring et al. 2013). With reference to the ESCAPE study the range of NO_2 exposures observed within the Swedish BAMSE cohort, the largest of the ESCAPE children's cohorts (44% of the total population) with the lowest exposures varied between 6.0-33.0 $\mu g/m^3$ (Gehring et al. 2103). Therefore, all of the annual NO_2 exposures attributions within this cohort were below the range observed in London: 31.2-98.9 $\mu g/m^3$. The higher concentrations observed in the current study, both reflect the more urban nature of the London cohort, but also the increase concentrations of NO_2 and NOx observed within European cities since the mid 2000's, due to the increased dieselization (Cames and Helmers 2013).

In the present study, I also examined the extent to which the lung function changes observed upon exposure to the urban air shed were modified by genetic variation in genes related to the metabolism of PM components. Environmental pollutants have been shown to be potent inducers of the expression of phase I and phase II metabolizing enzymes (Fujii-Kuriyama and Mimura 2005; Köhle and Bock 2007), with the metabolism PAHs, particularly from diesel exhaust particles inducing oxidative stress (Ma and Lu 2007). The AhR upon activation by PAHs, is translocated to the nucleus where it drives the transcription of numerous detoxification enzymes including NQO1, the GSTs and CYP1A1. These responses can then be further augmented by the imposition of oxidative stress, which also drives NQO1 and GST expression through the activation of the transcription factor Nuclear factor (erythroid-derived 2)-like 2 (Nrf2), which orchestrates the adaptive response of cells to various stressors (Cho et al. 2006). As there is considerable inter-individual variation in lung function, symptomatic and inflammatory responses to air pollutant exposures it has been argued that there is likely a genetic basis to individual sensitivity (Kleeberger et al. 2005; Cho & Kleeberger 2010). Two panels of SNPs were selected for these analysis, the first, designated the exploratory set examining SNPs in genes for which there was a previous literature demonstrating a potential influence on lung cancer, or lung function and symptom responses to air pollutants (Romieu et al. 2006; Wenten et al. 2009; Castro-Giner et al. 2009; Ren et al. 2010; Timofeeva et al. 2010); and the second exploratory set focused on genetic variants in AhR and CYP1A1, examining SNPs not previously associated with lung function/symptom interactions with ambient air pollution.

Of the selected candidate genes in the confirmatory panel, GSTM1 is the most widely studied. GSTM1 is located on chromosome 1p13.1, within the highly polymorphic GSTmu1 gene cluster (Pearson et al. 1993). Two major polymorphisms exist in GSTM1, GSTM1*O (a deletion allele), with the homologous allele (GSTM1-null) expressing no protein and GSTM1*A / GSTM1*B, with the later two alleles varying by a single base in exon 7, giving rise to active homo and heterodimeric enzymes (Strange et al. 2001). GSTM1 is predominately found in the liver, but is also expressed in the lung with the highest concentration in the major conducting airways

(Cantlay et al. 1994). The null genotype has been associated with enhanced airway inflammation in response to a diverse range of inhaled toxins: ozone (Alexis et al. 2009), endotoxin (Dillion et al. 2011), environmental tobacco smoke (Gilliland et al. 2006) and diesel exhaust particles (Gilliland et al. 2004). Numerous studies have also associated GSTM1 null with an increased risk of cancer (Carlsten et al. 2008; Hou et al. 2000), increased asthma susceptibility and reduced lung function (Gilliland et al. 2002; Tamer et al. 2004; Hanene et al. 2007), though the association of GSTM1 genotypes and asthma remains contentious (Reddy et al. 2010; Minelli et al. 2010). There is also some evidence that genetic variants in GSTM1 may be related to the rate of lung function decline in patients with chronic obstructive pulmonary disease, either singularly (Lakhdar et al. 2011), or in combination with other genetic polymorphisms (Chen et al. 2004). The potential role of GSTM1 in modulating pulmonary responses to pollutants was illustrated in recent in vitro work where the production of inflammatory mediators by primary bronchial epithelial cells challenged with diesel exhaust particles was shown to be enhanced by GSTM1 knockdown (Wu et al. 2012). Similarly, ozone-induced pulmonary inflammation has been shown to be augmented under controlled exposure conditions in subjects with the null genotype (Alexis et al. 2009). Individuals with the null genotype have been shown to have enhanced allergic responses to diesel exhaust particles (Gilliland et al. 2004).

In the present study I employed the tagging SNP rs366631 as a GSTM1 deletion marker, based on the work of Huang et al. (2009) demonstrating an association between this SNP and GSTM1 expression. I found no evidence that the GSTM1 null genotype was associated with basal lung function after adjustment for the available covariants, or the impacts of acute to chronic air pollutant exposures on lung function. This is broadly in line with the literature, where only a single published study has shown an interaction with the null genotype alone, reporting heightened small airway responsiveness (decreased FEF_{25-75}) and increased respiratory symptoms in asthmatic children (n=151-158) exposed to high ambient ozone exposures (Romieu et al. 2004, 2006). More commonly associations between the null genotype and pulmonary responsiveness have been reported as joint interactions with polymorphisms in other genes, predominately

NQO1 (Chen et al. 2007; Bergamaschi et al. 2001). Both the studies by Bergamaschi et al. (2001) and Chen et al. (2007) addressed lung function responses in adults exposed to ambient ozone, with the former showing a combined effect of NQO1 wild type (Pro187Pro) and GSTM1 null on airway function in cyclists before and after rides (>80ppb, n=24); and the later, a relationship between life time exposure to ozone and decreased FEF_{25-75}, for the same combined genotype, though only in female subjects (n=210). Whilst in the present study I also examined rs1800566 (Pro187Ser) there was no sufficient statistical power to examine genotype interactions. This is an important consideration as in the above studies there was no attempt to control for multiple comparisons made, in what were very small numbers of subjects. The results of the quoted interactions should therefore be interpreted with caution and overall one would have to conclude that the evidence that the GSTM1 null genotype is associated with poorer respiratory outcomes upon acute or chronic air pollutant exposures is weak, or at least unconfirmed.

GSTP1 (rs1695, Ile105Val, located at 11q13) is the predominate glutathione S transferase in the human airway (Fryer et al. 1986), with the GSTP1 valine allele (Val/Val), associated with reduced enzyme activity (Watson et al. 1998), reduced lung function (Caroll et al. 2005; Lee et al. 2005; Chen et al. 2007), increased susceptibility to tobacco smoke (Gilliland et al. 2006) and increased risk of early asthma (Melen et al. 2008; Tamer et al. 2004). In the present study we observed a valine allele frequency of 34.8% across the total population, consistent with previous reports (Kellen et al. 2007), with a particularly high frequency in the black population (44.5%). The majority of published studies looking at interactions between air pollutant exposures and GSTP1 polymorphisms have shown some evidence of effect modification, either alone or in combination with other SNPs, though the risk allele varies between studies (London and Romieu 2009; Minelli et al. 2010). Melen et al. (2008) reported increased allergic sensitization in children with the GSTP1 AG+GG genotypes in relation to their early life NOx exposures, whilst asthmatic children with the GG genotype (valine allele) in Mexico City experience enhanced ozone-related respiratory symptoms, though no evidence of enhanced FEV_1 or FVC responses (Romieu et al. 2006). A recent study

examining small cohort of South African children (9-11 years) demonstrated greater pollution-associated (PM_{10} and SO_2) increases in FEV_1 intra-day variability for the GSTP1 AG+GG genotype (Reddy et al. 2012). In contrast, in the present study, we observed no modifying effect of the polymorphisms in GSTP1 (rs1695 and rs749174) on the observed relationship between lung function and modeled pollutant exposures.

The final gene considered in the exploratory panel was NAD(P)H quinone oxidoreductase 1, which is involved in the detoxification of quinone compounds, either inhaled as a component of ambient PM (Baeza-Squiban et al. 1999), or formed as a consequence of PAH metabolism in vivo (Zhang et al. 2012). Of the polymorphisms considered, rs1800566 (Pro187Ser) is the most widely studied, with the serine allele associated with reduced (heterozygotes), or no (homozygotes) enzyme activity (Siegel et al. 1999). The combined GSTM1-NQO1 (Pro187Ser) genotype has been investigated with respect to ozone exposures in children where it has been associated with increased exhaled breath condensate (EBC) biomarkers of inflammation and oxidative stress (Corradi et al. 2002), increased pulmonary permeability (Bergamaschi et al. 2001), and decreased $FEF_{25-75\%}$ (Chen et al. 2007). In addition, genetic variants in NQO1 (rs10517, rs1800566, and rs2917666) have been shown to have increased asthma susceptibility in relation to their NO_2, exposure (Castro-Giner et al. 2009). In the present study, we observed no impact of 4 NQO1 SNPs on the association between air pollutant exposures, over any interval and changes in lung volume. It should be stated that where these, or similar associations have been examined in the previously quoted literature, no such associations have been reported.

Genetic variants in CYP1A1 and AhR genes were also investigated for their potential influence on the association between traffic-related air pollution and lung function. This represents the first time variants in these genes have been related to lung function air pollution interactions in the general population. In this exploratory analysis, 9 SNPs were examined, including 5 SNPs of AhR (rs2074113, rs2066853, rs17722841, rs17779352 and rs2282885) and 4 SNPs of CYP1A1 gene (rs2606345, rs1799814, rs17861115 and rs2198843). The aryl hydrocarbon receptor (AhR) is a member of the basic helix-loop-helix/PAS transcription factors family, located on chromosome 7,

known to be involved in chemical carcinogenesis. It can be activated to its DNA binding form by several exogenous xenobiotic ligands such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and importantly in terms of its likely involvement in modifying air pollutant effects, polycyclic aromatic hydrocarbons (PAHs). The activation of the AhR results in induction of CYP450 1A1, 1A2 and 1B1 enzymes (Denison and Nagy 2003). A number of genetic variants have been identified in the AhR (Harper et al. 2002), mainly present in exon 10, a region encoding most of the transactivation domains of the receptor, which regulates the expression of other genes, in particular CYP1A1 (Whitelaw et al. 1994). The most studied SNP located at exon 10 is Arg554Lys (rs2066853) was first identified in the Japanese population (Kawajiri et al. 1995). The same group has also discovered Asn44 (132T/C, rs17779352) located at exon 2. Subsequently, Smart and Daly (2000) identified SNPs at Val570Ile (1708G/A, rs4986826) and Pro517Ser (1549C/T, rs72552768) also within exon 10. Several studies have also reported polymorphisms at intronic regions such as rs2074113 (Chen et al. 2009; Kim et al. 2007), rs17722841 (Ng et al. 2010) and rs2282885 (Bin et al. 2008; Ng et al. 2010). There is growing evidence indicating that genetic polymorphisms in AhR cause significant changes in the response to xenobiotics (Maier et al. 1998), their toxicity in animal models (Nebert 1989; Okey et al. 1989) and increased cancer susceptibility in humans (Berwick et al. 2004; Gu et al. 2012; Long et al. 2006; Ng et al. 2010; Zhang et al. 2009). SNPs in the AhR (rs10250822 and rs2282885) have previously been shown to be associated with the increased excretion of urinary 1-hydroxypyrene in PAH exposed workers (Bin et al. 2008), though this association is yet to be observed in urban populations despite evidence of increased concentrations of this urinary metabolite (Lee et al. 2009). In the present study, no modifying effect of any of the studied Ahr SNPs was identified on the association between air pollutant exposure and lung function.

Cytochrome P450 1A1 (CYP1A1) is expressed in extra-hepatic tissues including the lung, where it is known to play an important role in PAHs metabolism and via AhR induction (Nebert & Ross 1996). There are a number of CYP1A1 polymorphisms, most notably the CYP1A1 Ile462Val (m2) allele, which has been shown to double the

activity of the microsomal enzyme in the heme-binding domain. In addition, mutations in CYP1A1 MspI (m1) have been shown to significantly increase catalytic activity in the Caucasian population (Landi et al. 1994). Polymorphisms in CYP1A1 have been investigated due to their important role in xenobiotic metabolism in relation to respiratory disease (Muñoz et al. 2012) and symptoms (Chen et al. 2011) associated with environmental tobacco smoke (ETS) exposure. However a recent genome wide association study failed to highlight CYP1A1 as a gene related to lung function risk in relation to cigarette smoking (Curjuric et al. 2012). Whilst several studies have investigated the influence of CYP1A1 polymorphisms on the relationship between air pollutant exposure and the generation of oxidative stress or traffic exposure biomarkers (Apostoli et al. 2003; Binkova et al. 2007; Novotna et al. 2007; Topinka et al. 2007; Adonis et al. 2003; Apostoli et al. 2003), this is the first investigation of their potential influence on pollutant and lung function interaction.

In the present study, genetic variation in the CYP1A1 gene was found to significantly modify the acute to sub-chronic effect of air pollution on lung volume, with rs2198843 (CYP1A1), being protective against and rs2606345 (CYP1A1) augmenting the reduction in FVC. Previous studies have shown that the variant allele 'T' of CYP1A1 rs2606345 (G/T) is significantly associated with the risk of smoking induced lung cancer (Canova et al. 2010; Rotunno et al. 2009). When these analyses were performed separately for each of the major ethnic groups it was apparent that these protective and sensitizing effects of the C and T alleles for rs2198843 and rs2606345, respectively, were seen only in the Asian and black populations. Thus the examined polymorphisms did not explain the larger negative association between annual pollutant exposures to NOx and NO_2 and FVC, observed in the white children.

There are notable differences between the current study and all of the previously cited literature. First, this is the only study examining the interaction of urban air pollution on lung function where measurements have been obtained post-bronchodilator. This is a significant issue, as this removes confounding due to undiagnosed or poorly managed asthma, which is high in this population (as reviewed in Chapter 3). Our study population is also multi-ethnic and lives in one of the most

deprived and polluted areas in the UK and perhaps most importantly this study reflects the contemporary urban air shed. A review of the previous literature, even that published within the last two years reveals that most of the cohorts investigated have linked air pollutant exposures to lung function measurement made in the 1990's to the early to mid-2000's (Gruzieva et al. 2013 [years covered 1994-2008]; Schultz et al. 2012 [1994-2001]; Gao et al. 2013 [1996-2003]; Nordling et al. 2008 [1998-2000]; Gehring et al. 2013 [2000-2007]; Eenhuizen et al. 2013 [2000-2001]; Hoek et al. 2012 [1988-1999]; Roy et al. 2012 [1993-1996]). Consequently, these studies largely report on associations based on historic air pollution and do not reflect the dramatic changes that have occurred in many European cities since the rapid dieselization of the fleet, and the introduction of particle traps. Globally, diesel cars have increased their share of the car market worldwide, with much of this growth in Europe, where more than one out of every two new cars is a diesel (Cames and Helmers 2013). A further important innovation in the present study was the use of scaling factors applied to the separate annual dispersion models to obtain spatially resolved estimates of shorter term exposures. As controlled and real world exposures to diesel derived aerosols have been shown to elicit rapid lung function decrements (McCreanor et al. 2007; Stenfors et al. 2004), it was therefore necessary to isolate acute from long term impacts on the measured lung function parameters. By employing the NOWCAST scaling factors, we were able to examine the short term effects of exposure on the same spatial domain as the annual exposures, which is key innovation in the present study. In my analyses, I examined three exposure intervals: acute (24h pre-visit), sub-chronic (7 day average pre-visit) and annual attributions. These intervals were selected based on known pathological mechanisms of inhaled xenobiotics, based on acute airway irritation (McCreanor et al. 2007; Stenfors et al. 2004), sub-chronic tissue adaptation and sensitization (Maes et al. 2010; Porter et al. 2007) and chronic airway remodelling (Churg et al. 2003; Novaes et al. 2007). Complex lag structures were not examined to prevent a high statistical penalty in relation to testing multiple exposure intervals. It is notable, that whilst many studies have taken considerable care to control for multiple comparisons in gene-environment interaction studies, exposure attributions have seldom been treated in a similar fashion (Baja et al. 2013; Ren et al. 2010).

There are a number of limitations with the present study. First, and perhaps most importantly, it was not possible due to the high mobility of children between schools in Hackney and Tower Hamlets to examine the impact of the introduction of the LEZ in a longitudinal cohort. Rather, we were reliant on separate cross sectional panels of children within each study year. Similar to the approach adopted by Schultz et al (2012) in the BAMSE cohort we attempted to dissect out the contribution of long versus short term exposures on lung function to infer information on lung growth, but it is important to state that even with our refined modeling approaches this is an imperfect solution. In terms of evaluating the impact of the LEZ itself on children lung function we were again challenged, both by the lack of a control population outside the zone, but also by the lack of evidence of a clear air pollution signal in NO_2 and PM_{10} related to the schemes implementation. Nevertheless the measured association between FVC and NOx, does suggest that effective reduction of this pollutant, or targeting it is predominate source (traffic within London) would achieve an important health dividend.

In summary, I observed significant negative associations ($P<0.05$) between FVC and each of the examined pollutants: using NO_2 as an illustration: -3.0 (-5.0-0.0) mL per $\mu g/m^3$ of pollutant, based only on the long term exposure attributions in a population of 8-9 year old children only, 85% of whom lived at a locations failing meet the European limit value for this pollutant. To place this in context, within the range of modelled annual NO_2 exposure attributions for the children in this study (30.9-98.9 $\mu g/m^3$, **Table 3.6**), assuming a mean FVC of 1.91L for the whole population, this would equate to losses varying between 92.7-296.7 mL, 4.85-15.5% of average FVC, assuming a linear dose function to zero and the absence of a plateu at the higher concentrations. More informatively, as a significant proportion of the studied children lived in areas exceeding the annual limit valued for NO_2 (40 $\mu g/m^3$), this would mean the at least 85% of the children in the cohort would have lost in excess of 120 ml (6.28%) of FVC, representing a large population with a clinically significant decrement.

Polymorphisms in GSTP1 and possession of the GSTM1 null genotype did not modify the underlying relationship between the modeled exposures and FVC, but clear

effect modification of the pollutant, FVC association was observed with polymorphisms in CYP1A1 with possession of the C allele being protective for rs2198843, whilst possession of the T allele for rs2606345 augmented the negative association. These data therefore demonstrate that exposure of primary school aged children to urban pollution is associated with decreased lung volume, even within areas where exposure contrasts are limited.

Chapter 5

Examination of Biomarkers of Response (*Oxidative Injury*) and Exposure (*Urinary Metals*) to Traffic Pollutants in Children

5.1 INTRODUCTION

In urban environments high exposures to primary traffic emissions, particularly from diesel vehicles have been shown to have negative health impacts on allergic (Janssen et al. 2003) and respiratory endpoints (Peters et al. 1999), including lung growth (Gauderman et al. 2004, 2007; Roy et al. 2012b; Schultz et al. 2012). This evidence base has been reviewed extensively in the recent USA Health Effects Institute review on the health effects of traffic pollutants (HEI 2010) and has been supported by the data presented in chapters 3 and 4. Mechanistically, these symptomatic responses have been related to the capacity of inhaled particulate and gaseous pollutants to elicit oxidative injury to the lung, related to their intrinsic capacity to oxidise biomolecules (nitrogen dioxide), or in the case of particulate matter their content of organic (polyaromatic hydrocarbons and quinones) and inorganic redox catalysts (transition metals) (Ciccone et al. 1998; Gerde et al. 2001; Li et al. 2003a/b; Valavanidis et al. 2008). Thus biomarkers of oxidative injury, or perturbations in antioxidant defences have been proposed as informative biomarkers of the biologic impact of pollutants on exposed populations (Barbato et al. 2010; Møller and Loft 2010; Weimann et al. 2012). Two biomarkers of oxidative stress that have been widely employed in the literature are the nucleotide and lipid oxidation products 8-oxo-7,8-dihydro-2-deoxyguanosine (8-oxodG) and 8-isoprostane respectively (Valavanidis et al. 2009). Air pollution exposure has been associated with an increased urinary excretion of 8-isoprostane (Lai et al.

2012) and 8-oxodG (Avogbe et al. 2005; Barbato et al. 2010). Therefore, in the present study these two markers were quantified in the children's urine to examine whether elevated exposures to urban air pollution was associated with evidence of systemic oxidative stress.

As the concentration of oxidised biomolecules in the body reflects an imbalance between the generation of reactive oxygen species and the endogenous antioxidant defences, I also examined whether the concentration of these biomarkers were influenced by polymorphisms in a panel of genes related to antioxidant defence: The transcription factor Nuclear factor (erythroid-derived 2)-like 2 (Nrf2), which also plays a major role in the induction phase I xenobiotic enzymes (Osburn and Kensler 2008); the glutamate-cysteine ligase regulatory subunit (GCLM), which regulates intracellular glutathione concentrations (Breton et al. 2011) and extra-cellular superoxide dismutase (SOD3), which represents the major antioxidant enzyme at the air-lung interface of the lung (Li et al. 2004). A previous in vitro study has demonstrated that genetic disruption of Nrf2 is associated with increased oxidative stress and DNA lesions (Reddy et al. 2008). Genetic polymorphisms in the Nrf2, GCLM and SOD3 have recently been associated with cardiovascular risk, acute lung injury and reduced lung function (Nakamura et al. 2002; Marzec et al. 2007; Dahl et al. 2008; Siedlinski et al. 2009).

As a secondary exploration within this analysis, I also examined whether quantification of urinary metals could be employed as an exposure biomarker for traffic. Previous studies have employed urinary metals analysis to assess occupational or environmental exposures to aluminum (Riihimäki and Aitio 2012), arsenic (Saoudi et al. 2012), mercury (Zubero et al. 2010; Castaño et al. 2012), cadmium (Zubero et al. 2010; Castaño et al. 2012; Cerná et al. 2012), nickel (Aguilera et al. 2010; Huang et al. 2013), copper (Aguilera et al. 2010; Callan AC et al. 2012; Huang et al. 2013), lead (Zubero et al. 2010; Castaño et al. 2012) and chromium (Aguilera et al. 2010; Zubero et al. 2010). Therefore, I examined an expanded panel of 15 metals, selected to reflect different source contributions in ambient air (Thorpe and Harrison 2008). These have been summarised below in **Table 5.1**.

Table 5.1 Sources of the most abundant metals associated with ambient PM

Metal	Environmental Sources
Copper (Cu)	Copper emissions are from a variety of sources including copper processing, smelting, and other metal production. It is also a byproduct of municipal incineration and coal and oil combustion. In terms of traffic related sources, copper is a large component of brake wear.
Iron (Fe)	Iron emissions are primary sourced from abrasion emissions, such as non-exhaust traffic emissions and as a by-product of the iron and steel industry. It used to be a large component of fuel oil combustion, but its content in modern fuel has decreased significantly.
Manganese (Mn)	Most commonly enters the air through the production of ferromanganese compounds in blast furnaces. Also, found in welding rods, the incineration of manganese-containing products and is also used as a fuel additive.
Nickel (Ni)	The direct sources of nickel in the atmosphere are from nickel ore mining and smelting as well as secondary and co-product nickel recovery. There are also a number of indirect sources of nickel in the air including: coal and oil combustion, and petroleum processing. Small amounts can be attributed to municipal refuse and sewage incineration.
Chromium (Cr)	Direct sources of chromium are chromite ore refining, ferrochromium and refractory and chromium chemical production, steel production and leather tanning. Indirect sources are coal and oil combustion, cement production, cooling towers and municipal refuse and sewage incineration.
Antimony (Sb)	Antimony is used as part of various alloys in the production of many different metals, such as various semiconductor devices. It is also a common constitute of paint pigments as well as rubber compounding. As well as being released directly into the air through wear of various items which contain antimony, such as brake lining pads, antimony in the atmosphere is often released into the air from smelting plants.
Cadmium (Cd)	Pure cadmium is not found naturally in the environment; therefore, any component of cadmium in air particulates is sourced as a by-product of refining other metals, mostly zinc. It is also found in numerous fertilizers and can be produced as non-exhaust emission from traffic.
Arsenic (As)	Although there are a variety of sources of arsenic in the environment, the majority of the particulate matter of arsenic is sourced from combustion of various products including, fuel, various types of waster as well as from lime kilns and crematories.
Molybdenum (Mo)	Molybdenum in the environment can come from a variety sources including weathering, agriculture and industrial processes. However the largest component comes from the burning of fossil fuels.
Barium (Ba)	The major source of barium in the atmosphere is emitted from smoke emissions either from diesel engines or as by-products from various industrial processes such as mining and refining. Barium is also a byproduct of brake wear from various motor vehicles.
Vanadium (V)	The major sources of vanadium in the atmosphere include vanadium refining and alloy production, power plants and combustion of vanadium-rich oils.
Zinc (Zn)	Zinc is emitted into the atmosphere from refineries for which zinc is either the primary or the secondary product, or as a by-product of the production of other metals. It is also a product of non-exhaust traffic emissions, specifically tyre wear from motor vehicles.
Titanium (Ti)	Approximately 90% of the atmospheric emissions of titanium are not associated directly with the titanium industry. Instead, the major source of titanium in the air is from the combustion of coal, and fossil fuels.
Tungsten (W)	Tungsten is not found naturally as a pure metal; it has to be extracted from mineral ores to produce pure tungsten metal, alloys or compounds. It is released into the atmosphere as windblown dusts. In automobiles, it is sourced from abrasion at the road surface.
Tin (Sn)	Tin is released into the atmosphere from a variety of different sources. Tin is a common component of soil and therefore can be released into the air through weathering or agricultural activities. It is also released into the atmosphere from abrasion of items containing tin and in dusts from wind storms and roads.

Based on information reviewed in Copper and Harrison (2009); Godri et al. (2011); Harrison and Yin (2011); Thorpe and Harrison (2008), Kelly et al. (2011b).

Whilst the aim of this secondary study was to examine whether urinary metals could be used as informative traffic biomarkers, it was appreciated that diet would be a major confounding factor. Thus, it was important to understand the potential source of the studied metals in food. Important dietary sources of these metals are listed in the **Table 5.2.** Most of these metals are classified as essential or trace elements and are only found in small quantities in the body to maintain the homeostasis.

This study therefore represents one of the largest simultaneous examinations of urinary biomarkers of oxidative stress and traffic exposure performed on a school age urban population. It also represents the first time that the concentrations of these biomarkers have been related to modelled pollutant exposures at the residential address level, over a range of exposure windows, allowing the short – versus long-term impacts of air pollution on these markers to be investigated. In this analysis I hypothesised that children with high pollutant exposures would have elevated urinary concentrations of 8-isoprostane and 8-oxodG, associated with increased concentrations of urinary metals, related to both exhaust and non-exhaust traffic sources. I also examined whether polymorphisms in three key antioxidant genes resulted in increased concentrations of the oxidative biomarkers, reflecting an impaired capacity of the lung to adapt against pollutant induced oxidative stress.

Table 5.2 Potential dietary sources of the metals analyzed in this study

Metal	Dietary Sources
Copper (Cu)	Dietary sources include shellfish, whole grains and organ meats (kidney, liver), as well as dark leafy greens (Trumbo et al. 2001).
Iron (Fe)	Dietary iron sources can be divided into two separate forms, heme and nonheme. Heme iron is sourced from animal foods and non-heme iron comes from plant foods (Hurrell 1997).
Manganese (Mn)	Higher manganese exposure is normally seen in vegetarians as grains, beans and nuts have high manganese contents. There is also a higher exposure to manganese in people who drink large amounts of tea (Department of Health and Human Services 2008).
Nickel (Ni)	A large proportion of the dietary intake of nickel is due to leeching of nickel-containing alloys in food processing equipment. Certain foods such as oatmeal, dried beans, peas, nuts and soya products are all high in nickel content (Barceloux 1999; Arnich et al. 2012).
Chromium (Cr)	Although chromium is found in a variety of foods, it is often only in very small quantities. The foods with the highest concentrations are meat and whole grain products (Anderson 1992).
Antimony (Sb)	The most common dietary source are meats and fishes, but concentrations are very low (Arnich et al. 2012).
Cadmium (Cd)	The most common dietary sources of Cadmium are shellfish and organ meats as well as leafy vegetables and soya produces. Food is one of the most common sources of cadmium exposure in non-smokers (Robards and Worsfold 1991).
Arsenic (As)	Arsenic is commonly found in rice and rice products (Jackson et al. 2012; Arnich et al. 2012).
Molybdenum (Mo)	The foods with the highest concentration of molybdenum are normally animal items, specifically liver and kidneys (Schroeder et al. 1970).
Barium (Ba)	Dietary sources with the highest concentration of Barium include brazil nuts, seaweed and fish (Department of Health and Human Services, 2007).
Vanadium (V)	While most foods naturally contain low concentrations of vanadium, the highest concentrations can be found in seafood products (Badmaev et al. 1999).
Zinc (Zn)	Zinc is found in highest concentrations in meats, fish and poultry. However the largest dietary contributor to zinc intake is water and other beverages, especially if these beverages are stored in metal containers (Gibson 2012).
Titanium (Ti)	Amongst the sources contain titanium are chocolate, candies, chewing gums, and powdered sugar toppings (Weir et al. 2012).
Tungsten (W)	Limited information available, but dietary intake appears to be extremely low (Agency for Toxic Substances and Disease Registry (ATSDR) 2005).
Tin (Sn)	As tin is naturally found in soil, it can be present in low amount in vegetables and fruit and the most common source of tin is through consuming canned foods (Shimbo et al. 2013).

5.2 METHODS

5.2.1 Study Design and Population

Details of study design and volunteer population are presented in C**hapter 4, Table 4.1**. Subjects provided urine samples during the study visits, aliquots of which were then stored at -80^{o}C until required for analysis. Excess urine was discarded according to the regulations stipulated by the Human Tissue Authority.

5.2.2 Urinary Analysis

Details of creatinine and cotinine measurements are presented in **Section 2.3.5** (**Chapter 2**).

5.2.2.1 Determination of Urinary Biomarkers of Oxidative Stress

5.2.2.1.1 Analysis of 8-hydroxy-2'deoxyguanisine

The concentration of 8-hydroxy-2'-deoxyguanisine or 8-oxodG, a biomarker for oxidative DNA damage, was determined using a commercial competitive enzyme-linked immunosorbent assay (ELISA) (Oxis Health Products, Inc –Cat#:21026, Portland OR, USA). 50 µl sample or standard, plus 50 µl of monocolonal 8-oxodG antibody (diluted in phosphate buffered saline - PBS) were added to wells pre-coated with 8-oxodG and the plate incubated at 37^{o}C for an hour. To avoid edge effects, the outermost wells of the plate were not used. The incubation conditions creates a competitive environment between 8-oxodG derived from the sample with that pre-bound to the plate for monoclonal antibody binding. After the incubation period, the unbound antibodies were washed away and a secondary enzyme-labelled antibody added, diluted in PBS. Any unbound secondary HRP-conjugated antibody was then washed away after a second incubation at 37°C for 1 hour, prior to the addition of the chromogen (3,3',5,5'-

Tetramethylbenzidine diluted in hydrogen peroxide/citrate-phosphate buffered saline). After a further 15 mins incubation at room temperature in the dark, the colour development was stopped by the addition of 100 μL of the supplied stop solution (1M phosphoric acid). The absorbance of the plate was then measured at 450nm using a plate reader (SpectraMAX 190; Molecular Devices)**.** Concentrations were determined with reference to a 6 point standard curve (0.5 ng to 200 ng/ml) using a 4-parameter logistic curve fit. Final concentrations were normalised against urinary creatinine concentrations. A representative of 8-hydroxy-2'deoxyguanisine standard curve is illustrated in **Figure 5.1.**

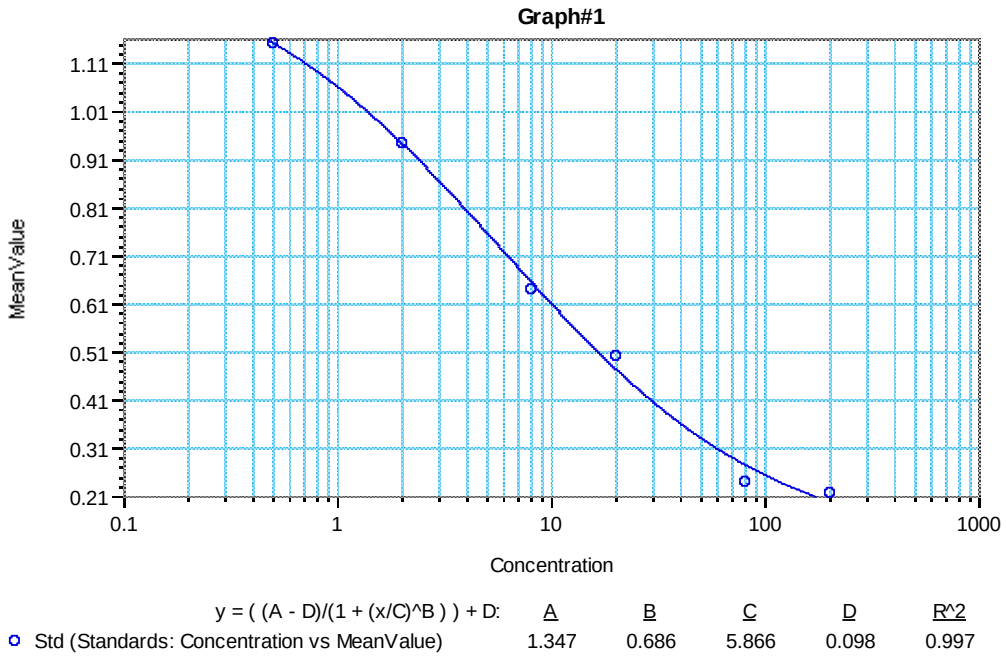


Figure 5.1 Representative of 8-oxodG assay standard curve. The y-axis represents the mean absorbance values at 450 nm, the x-axis the log of the standard concentration.

5.2.2.1.2 Analysis of 8-isoprostane

A commercially available competitive ELISA assay (Cayman Chemical) was used for the determination of 8-isoprostane concentrations in the urine samples. The

assay was based on the competition of 8-isoprostane within the sample with a 8-isoprostane-acetylcholinesterase (AChE) tracer for a limited number of 8-isoprostane-specific rabbit antisera binding sites within the plate wells. As the tracer concentration is held constant, whilst the 8-isoprostane varies between samples and standards, the amount of AChE that binds to the antisera is inversely proportional to the 8-isoprostane concentration in the samples. For the actual assay, 100 ul buffer was added to the non specific binding wells, whilst the maximum binding wells received 50 ul of the buffer, plus 50 ul of samples (urine samples were diluted 20-times prior to the assay) or standards (0.8-500 pg/ml). 50 ul of 8IsoP tracer and 50 ul of 8-isoprostane antiserum were then added to each well, with the exception of the total activity and blank wells after which the plate was incubated at 4^{o}C for 18 hours. After the incubation period the wells were emptied and washed five times with the supplied wash buffer, prior to the addition of 200 ul of Ellman's Reagent. Following further 90 mins incubation in the dark, the absorbance of the plate was measured at 412 nm using a plate reader (SpectraMAX 190; Molecular Devices). Quantification was achieved relative to the standard curve, with a detection limit of 2.7 pg/ml (**Figure 5.2**). The analysis was repeated when coefficient of variation (CV) was >10% for 8-isoprostane concentration of >32 ng/mmol and >20% for the concentration <32 ng/mmol. The expected 8-isoprostane concentration in urine was normalised against creatinine concentrations by converting the values from ng/mmol to mmol/ml.

5.2.2.2 Determination of Urinary Metals Concentrations

5.2.2.2.1 Acid Digests of Urine Samples

Metal free conical bottom centrifuge tubes of 15 ml capacity with screw cap necks were used for the urine digestions (Elkay Laboratory Products, Basingstoke, England – product No 2086-500). Prior to use, the tubes were thoroughly washed in two stages in order to reduce potential metal contamination. First, with Chelex-resin treated water (3 times), after which they were dried in a fume cupboard for 48 hours, followed

by washing with 3 ml of 6.5% nitric acid solution (10.8 ml of 60 % HNO_3 (Merck. Darmstadt, Germany – product code 1.01518.05800) and 89.2 ml of Chelex-100 resin treated water). 500 ul of urine was then added to 1500 ul of freshly prepared 6.5% Nitric acid solution in the presence of 20 ul internal standard Yttrium (1ppm Yttrium (Heeman Lab, Inc, Lowell, MA, USA) in 6.5% Nitric acid) and samples incubated at 90^{o}C for 1.5 hours in a thermostatically regulated water bath. After cooling overnight at room temperature, samples showing precipitation were centrifuged for 10 mins at 4000 rpm. Subsequently, the supernatants were transferred to new acid-prewashed tubes and stored at 4°C until analysed. All water used for tube washing and dilution of nitric acid was deionised and ultra-filtered using an Elga-stat filtration system, prior to Chelex-100 resin treatment: 3 g of Chelex-100 resin per 100 ml, with overnight stirring at 4^{o}C. The resin was subsequently removed by centrifugation (3,000 rpm for 15 mins, 4^{o}C). Prior to use the pH of the water-methanol solution was adjusted to neutrality using Chelex resin treated 1M HCl or 1M NaOH.

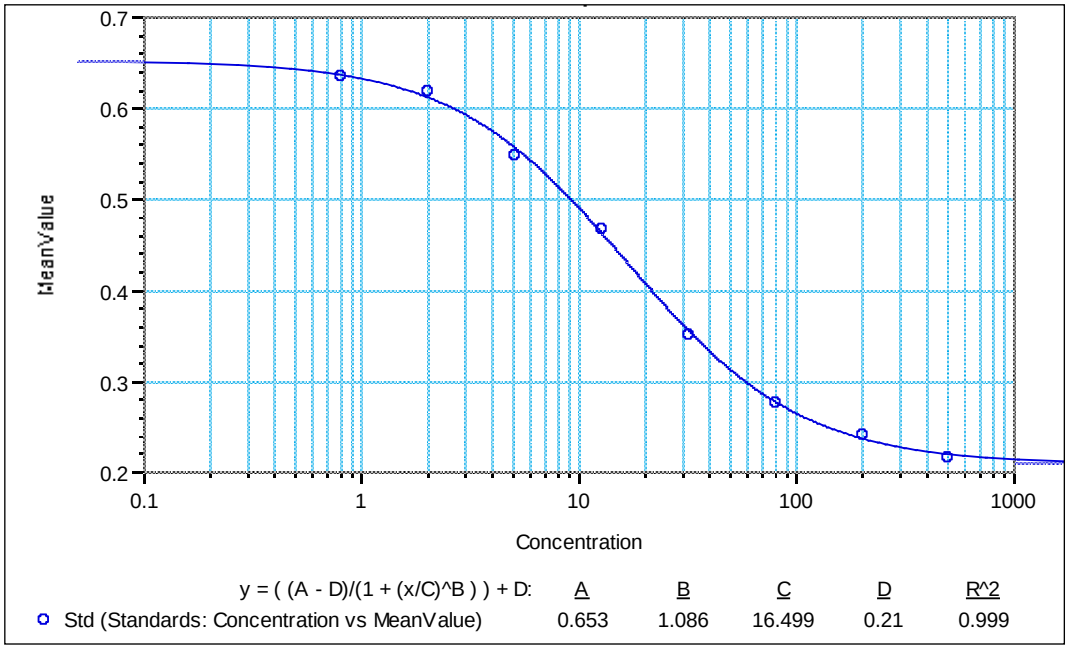


Figure 5.2 Representative of 8-isoprostane assay standard curve. Absorbance (A_{450} nm) was plotted against standard 8-isoprostane concentration. The sensitivity and minimum detection limit of this assay was defined as 2.7 pg/ml and the intra- and inter-assay variability range 12.6 - 34.8% and 10.5 - 39.1%, respectively.

5.2.2.2.2 ICP-MS Analysis

Samples were then transferred to the Mass Spectroscopy Unit at King's College London for the quantification of Cu (isotope 63; natural abundance 69.17%), Ba (137; 11.23%), Fe (56; 91.75%), Mn (55; 100%), Mo (95; 15.92%), Ni (60; 26.22%), Cr (52; 83.79%), V (51; 99.71%), Sb (121; 57.21%), Zn (66; 27.90%), Cd (111; 12.80%), As (75; 100%), Ti (47; 7.44%), W (184; 30.64%), Sn (118; 24.22%) and the internal standard Y (89; 100%) (Audi et al. 1995; Rosman and Taylor 1995) by ICP-MS using a ELAN DRC ICP-MS (MSF008). All analyses were performed by Dr A Cakebread. The potential $^{40}Ar^{16}O^{+}$ interference for the major isotope of iron (^{56}Fe) was removed using the dynamic reaction cell through its reaction with ammonia. The use of the DRC was also employed for the following elements: Cu (potential interference with $^{23}Na^{40}Ar^{+}$), Mn ($^{40}Ar^{14}N^{1}H^{+}$), Ni ($^{44}Ca^{16}O^{+}$), Cr ($^{35}Cl^{17}O^{+}$ and $^{40}Ar^{12}C^{+}$), As ($^{40}Ar^{35}Cl^{+}$), V ($^{35}Cl^{16}O^{+}$, $^{33}S^{18}O^{+}$) and Zn ($^{34}S^{16}O_2^{+}$, $^{50}Ti^{16}O^{+}$). Sn was also measured following reaction with ammonia in the DRC to remove the high background, probably reflecting isotopic interference from $^{12}C_4^{+}$. Elemental concentrations were determined with reference to a 7-point standard curve based on an ICP multi element standard solution VI CertiPUR® (Merck, Lot. No. OC529648). Antimony which is not present in this multi-elemental standard was analysed against its own standard curve (Antimony ICP standard, MERCK). All concentrations were corrected for the background elemental concentrations determined in the Chelex-100 resin treated water blank ran in parallel to each batch of samples.

5.2.3 DNA Sample Collection and SNP Genotyping

Details of DNA sample collection and the genotyping procedure are presented in **Section 2.4**. Briefly, saliva samples were collected from the children using the Oragene DNA kit OG-250 (DNA Genotek Inc, Canada) as a source of genetic material. The collection was carried out according to the manufacturer's instruction. Subsequently, genomic DNA was recovered from the saliva using ethanol precipitation technique. In total, 10 SNPs were investigated in three genes with functions related to protection

against or adaption to oxidative stress: Nrf2 (rs2364723, rs7557529, rs2001350), 2 SNPs of GCLM (rs2301022, rs3170633) and 5 SNPs of SOD3 gene (rs2284659, rs8192287, rs699473, rs13306703, rs8192288). Three Tag SNPs in Nrf2 (NFE2L2) were selected based on their reported asscociation with oxidant-related disease risks. SNP rs2364723 has been associated with lower basal FEV_1 in two Dutch cohorts, as well as with increased FEV_1 in smokers (Siedlinski et al., 2009). In addition, minor allele carriers of rs2364723 have been reported to have reduced risk of cardiovascular mortality in the same cohort (Figarska et al., 2014). This SNP is in almost complete linkage disequilibrium with the recently described promoter polymorphism rs35652124, suggesting the above assocation may reflect a functional impairment of Nrf2 signalling. Similarly, rs2001350 has been associated with the annual rate of FEV_1 decline, in a haplotype (containing the functional SNP, rs6721961) within the japanesse population (Masuko et al., 2011). This SNP has also been associated with Parkinson disease risk, as has rs7557529 (5238G>A) (von Otter et al., 2010). The two candidate SNPs in GCLM (A/G), rs3170633 (A/G), were based on those examined by Ren et al (2010), which had previously been associated with decreased glutathione concentrations in schizophrenics, as well as increased risk of this condition (Tosic et al., 2006). For SOD3 we examined rs8192287 (within the noncoding, 5′ untranslated region in exon 1) and rs8192288 (first interon) based on the previous observation of Dahl et al (2008) that individuals homozygous for the minor alleles had reduced FVC in two large, population-based studies. The remaining Tagging SNPs, rs2284659, rs699473 and rs13306703 were selected based on the analysis of Arcaroli et al (2009) to caputure haplotypes with a frequency greater than 2% in the European American population.

All selected SNPs displayed a >5% minor allele frequency, determined from the published web based data sets (1000 genome data set: integrated phase 1, version 3, March 2012; HapMap (Thorisson et al., 2005)). Details of linkage disequilibrium are presented in **Appendix J**, **Table J2**.

Subsequently, all the subjects were typed for the SNPs by GoldenGate genotyping assay on Illumina BeadXpress platform (Illumina Inc., San Diego, USA) and the genotype data was analyzed for quality control using the BeadStudio software.

Details of Hardy-Weinberg equilibrium p-values, genotype and allele frequencies for all SNPs are presented in **Appendix K**, **tables K1** and **K2**. All SNP markers included in the analysis conformed to HWE (p>0.05) except for SNP SOD3 rs2284659, therefore this SNP was excluded from further analysis. Overall genotyping success rate was >99.6% (**Table G1** in **Appendix G**).

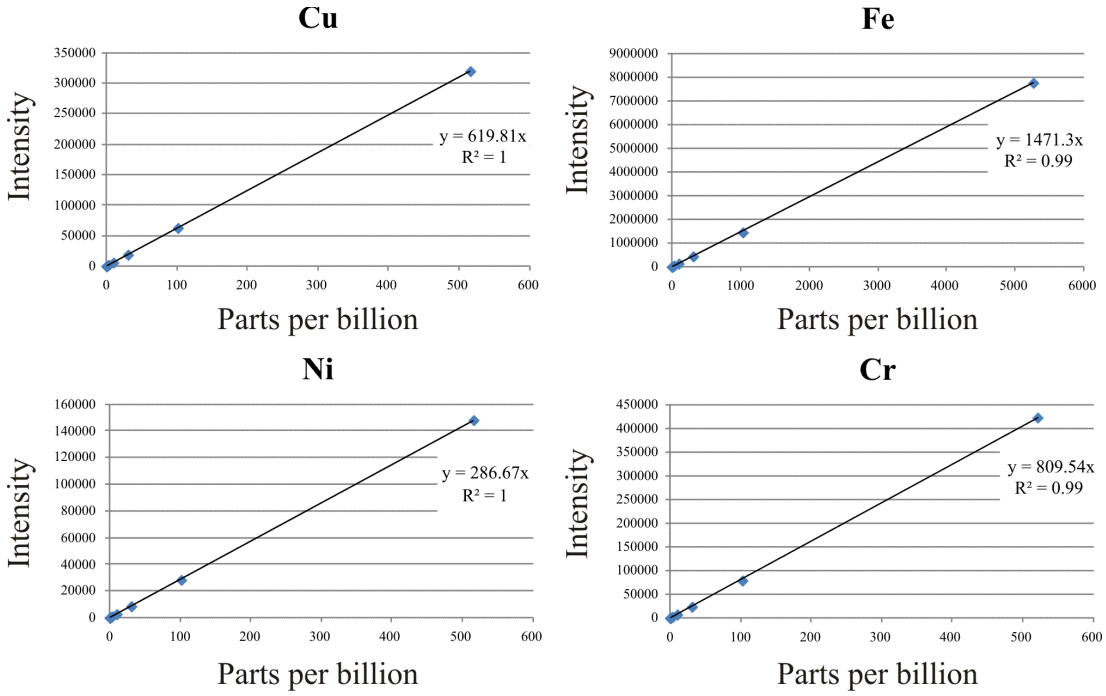


Figure 5.3 Representative elemental standard curves for Cu, Fe, Ni and Cr.

5.2.4 Assessing Air Pollution Exposures

The generation of modelled annual and temporally adjusted PM_{10}, $PM_{2.5}$, NO_2 and NOx concentrations are presented in **Chapter 3**, **Section 3.2.4**.

5.2.5 Statistical Analysis

All urinary biomarkers displayed highly skewed distributions and not normally distributed, based on the Shipiro-Wilks test and examination of quantile plots. For the descriptive analysis therefore these data are presented as median values with inter-quartile range (25^{th}-75^{th} percentiles). A one way ANOVA using Kruskal-Wallis rank test was used to investigate any differences between groups, defined by school location or ethnicity, with the Mann-Whitney U-test employed for post-hoc testing. The Spearman Rank Order Correlation test was used to assess association between the measured biomarkers and the modelled pollutant exposures.

A more detailed examination of the data was performed using a linear mixed model approach with random effect for school; to examine the effects of individual modeled pollutant exposures on natural log transformed urinary oxidative stress and metal outcomes. Due to the strongly right skewed urinary metal distributions, a generalized linear model with gamma errors was applied. Negative values (generated after subtraction of the blank digest concentrations) and extreme outliers were excluded from the analysis. To address the influence of bias introduced by the study population a number of baseline characteristics were selected prior to their inclusion in the models: age, sex, height, body mass index (BMI), self-reported ethnicity (Asian, Black, White, Others) and exposure to tobacco as measured by urinary cotinine:creatinine ratio >30ng/ml. To assess the assumption of no change over time, study year was also included as a covariate in the models. All results were back transformed and can be interpreted as a % change.

Of the 90 single nucleotide polymorphisms (SNPs) studied, 10 were selected a priori in three genes related to the regulation of oxidative stress (Nrf2, GCLM and SOD3) for the purpose of examining their influence on the relationship between concentrations of urinary biomarkers of oxidative stress and modeled pollutant exposures. Each genotype was recorded as a categorical variable representing wild type, heterozygous or homozygous genotype. In order to distinguish any direct effects of genotype on oxidative stress a crude and an adjusted analysis was first performed for

each SNP. The adjusted analysis included the same covariates as described above except air pollution exposures. To then test for effect modification of the SNPs, interaction terms of each categorical SNP and each continuous air pollution exposure attribution were added to the linear mixed models. For each SNP, the genotype with largest frequency for the total population (see **Table K2** in **Appendix K** for genotype and allele frequency for each SNP) was deemed to be the wild type and used as reference category for the comparison between the three categories. For cases where the effect modification was significant for only one comparison but not the other, a global test was performed to assess influence of the SNP overall. Additionally, we performed all analysis combining heterozygous and homozygous into a single category, which increased the power to detect effects and also confirmed results from the global analysis.

For each model, residuals were assessed for normality using quantile plots of the variable against the normal distribution and checked for heteroscedasticity by plotting residuals against the fitted values on the school level. Multiple testing was performed on all significant and non-significant p-values using a step-up false discovery rate (FDR) procedure as originally proposed by Simes (1986) and later justified by Benjamini & Yekutieli (2001). As described by both groups, controlling the false discovery rate is less conservative than using the family-wise error rate and therefore results in more differences being significant at the cost of some confidence in the results. The Simes' FDR procedure calculates a corrected critical p-value for each observation, so that an individual null hypothesis is considered to be acceptable if its p-value is greater than its corresponding corrected critical p-value. Models were examined on the 5% significance level for two-sided tests of individual associations and on the 10% significance level for interaction terms of effect modification. All statistical analyses were performed using Stata 10.1 (StataCorp, College Station, TX, USA) and SPSS, version 17.0 (SPSS Statistics for Windows, Version 17.0. SPSS Inc., Chicago, USA).

5.3 RESULTS

5.3.1 Sociodemographic Characteristics of Study Population

In total 1019 children participated in the study between Nov 2008 and April 2011, with 974 providing a spot urine sample for biomarker analysis. The demographics related to this restricted group are summarized in **Table 5.3**, together with information on the pollutant attributions for study years 1 – 3 combined, based on annual, 7 day and 24 hour exposures modeled on the subject's residential address, as previously described **Chapter 3**. In addition, **Table 5.3** also provides raw descriptive statistics on the urinary biomarker concentrations. For the 8-isoprostane determinations urine was available from 972 subjects, for 8-oxodG, 922, and for the metal analysis 883, reflecting sample availability. These data demonstrate that the biomarker concentrations, particularly of the selected metals (Ni and Cu) are highly skewed.

5.3.2 Descriptive Analysis of Oxidative Stress Biomarkers

Both 8-isoprostane and 8-oxodG concentrations varied significantly ($P<0.001$) across the 23 school sites, for the children recruited over the first 3 years of the study – **Figure 5.4**, **Panels A** and **B**. The two biomarkers of oxidative stress were weakly, but significantly associated (p=0.150, $P<0.001$ - Spearman Rank Order correlation), Figure **5.4**, **Panel C**. Children from Asian backgrounds showed significantly higher values for 8-isoprostane, with Black children having the lowest values (**Figure 5.4**, **Panel D**). Similar results were obtained for 8-oxodG. No evidence was apparent for sex differences in the concentration of these biomarkers (**Figure 5.4**, **Panel E).**

Table 5.3 Baseline characteristics for subjects in the London Low Emission Zone school study from who spot urine samples were collected

Characteristics	Subjects (n=974)			
	n	% or Mean	Min	Max
Age, mean (SD)	974	8.8 (0.3)	8.1	9.7
Sex				
Girls	466	47.8		
Boys	508	52.2		
Height (cm), mean (SD)	965	133.8 (6.7)	111.7	156.2
Weight (kg), mean (SD)	965	32.5 (7.9)	18.8	65
BMI, mean (SD)	965	18.0 (3.2)	11.8	31.3
Overweight (define as BMI 91st centile UK90 growth reference)	86/965	8.9		
Ethnicity				
Asian	345	35.4		
Black	245	25.3		
White	269	27.6		
Mixed	115	11.8		
ǂ Household second hand smoke exposure (cotinine)	229/971	23.6		
ǂ Social deprivation (IMD), mean (SD)	972	45.4 (11.1)	5.6	66.2
8-isoprostane (ng/mg, corrected for creatinine), (log) mean (SD)	972	0.1 (0.4)	-1.3	2.1
8oxodG (nmol/mmol, corrected for creatinine), (log) mean (SD)	922	1.9 (0.7)	-1.4	5.5
Copper (nm/mmol, corrected for creatinine), (log) mean (SD)	953	4.4 (0.7)	1.9	11.5
Nickel (nm/mmol, corrected for creatinine), (log) mean (SD)	882	2.5 (0.8)	-0.008	6.5
8-isoprostane (ng/mg, corrected for creatinine), median	972	1.06	0.28	8.43
8-oxodG (nmol/mmol, corrected for creatinine), median	922	6.72	0.24	251.86
Copper (nm/mmol, corrected for creatinine), median	953	72.64	6.55	4814.71
Nickel (nm/mmol, corrected for creatinine), median	882	10.69	0.99	653.81
Annual NOx 20m ($\mu g/m^3$), mean (SD)	971	75.8 (14.7)	47.5	235.4
Annual NO_2 20m($\mu g/m^3$), mean (SD)	971	43.5 (5.5)	31.2	98.9
Annual PM_{10} 20m($\mu g/m^3$), mean (SD)	971	23.4 (1.4)	20.5	33.2
Annual $PM_{2.5}$ 20m($\mu g/m^3$), mean (SD)	971	13.7 (0.8)	12	18
7 days NOx 20m, ($\mu g/m^3$), mean (SD)	971	46.5 (16.6)	16.9	181.3
7 days NO_2 20m, ($\mu g/m^3$), mean (SD)	971	25.0 (6.0)	12.2	67.3
7 days PM_{10} 20m, ($\mu g/m^3$), mean (SD)	971	22.7 (5.5)	13.6	60
7 days $PM_{2.5}$ 20m, ($\mu g/m^3$), mean (SD)	971	13.6 (3.7)	8.2	35.3
24 hour NOx 20m, ($\mu g/m^3$), mean (SD)	971	58.1 (43.3)	9.5	396.2
24 hour NO_2 20m, ($\mu g/m^3$), mean (SD)	971	27.8 (10.8)	8.1	75.5
24 hourPM_{10} 20m, ($\mu g/m^3$), mean (SD)	971	24.3 (9.4)	8.7	66
24 hour $PM_{2.5}$ 20m, ($\mu g/m^3$), mean (SD)	971	14.9 (6.5)	6.3	44.9

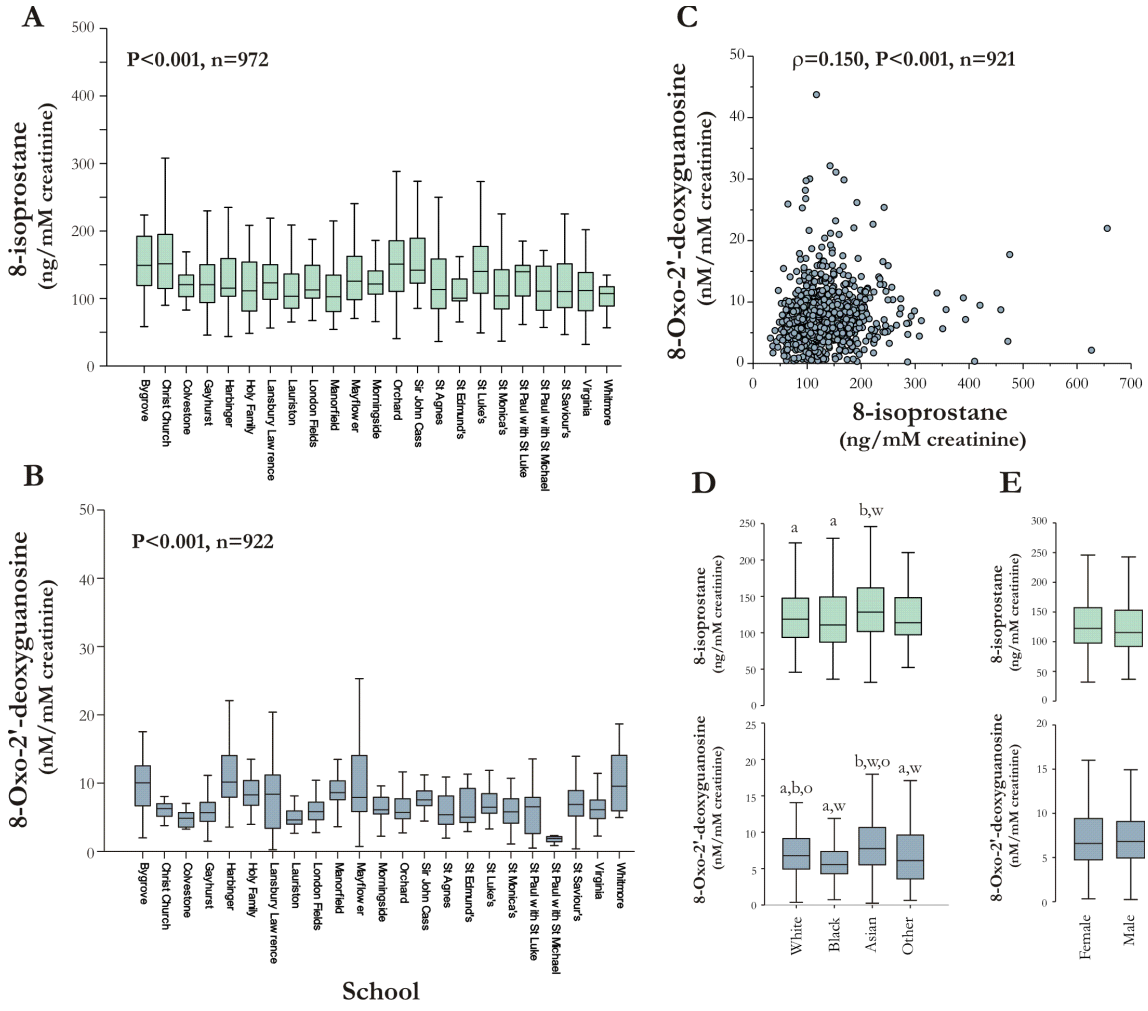


Figure 5.4 Descriptive statistics for the urinary 8-isoprostane and 8-Oxo-2'-deoxyguanosine. Data are illustrated as box plots, with the central line illustrating the median, the lower and upper boundaries of the box the 25th and 75th percentiles and the whiskers the 95% confidence intervals. **Panel**s **A** and **B** show the concentrations of two biomarkers (8-isoprostane and 8-OxodG) in all subjects for years 1-3 grouped according to the school locations. **Panel C** depicts the correlation between 8-isoprostane and 8-OxodG concentrations. **Panels D** and **E** represent 8-isoprostane and 8-OxodG concentrations segregated by ethnicity and sex. For **Panel D** 'a' represents a significant statistical difference (P<0.05) between the various populations compared to the Asian population; 'b' compared with the Black population; 'w' related to the White population and 'o' shows comparison with the subjects of 'Others' group.

The exposure of children to the environmental tobacco smoke (ETS) at home was also investigated. **Figure 5.5** shows the comparative concentrations of 8-isoprostane and 8-oxodG in children exposed to ETS with the non-ETS exposed

children, based on their measured cotinine / creatinine ratio – as described in **section 3.2.3.2**. The results indicated that the children exposed to ETS showed significantly higher concentrations of both 8-isoprostane and 8-oxodG: for 8-IsoP, 118.59 (92.76 - 153.2) (n=750), versus 125.47ng/mM (101.43 - 160.42) ng/mM creatinine (n=219), P<0.05; and for 8-oxodG, 6.51 (4.61 - 9.12) nM/mM creatine versus (5.52 - 9.93) nM/mM creatinine (p<0.001).

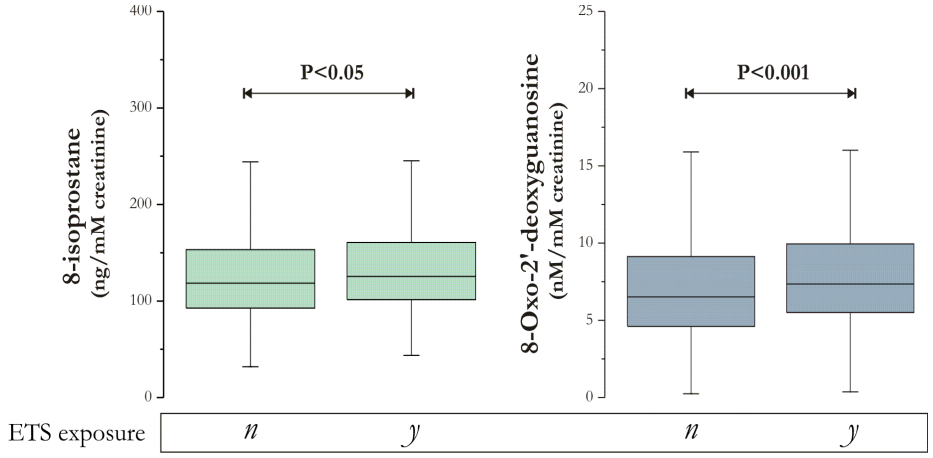


Figure 5.5 Effect of ETS exposures on the 8-isoprostane and 8-oxodG concentrations. Notes: *n*=no, not exposed to ETS; *y*=yes, exposed to ETS.

The associations between the urinary oxidative stress markers and modeled exposures to NOx, NO_2, PM_{10} and $PM_{2.5}$ are illustrated in **Figure 5.6** – **5.9**. Overall, there was evidence of weak but significant negative associations between acute 24 hour exposures to NOx and NO_2 and urinary 8-isoP concentrations – **Figure 5.6** and **5.7**. No significant associations were seen with acute to chronic exposure attributions for PM_{10} or $PM_{2.5}$. The pattern with 8-oxodG was somewhat different with weakly positive associations with 24 hour NOx, NO_2 (**Figure 5.6** and **5.7**) and PM_{10} (**Figure 5.8**) exposure attributions.

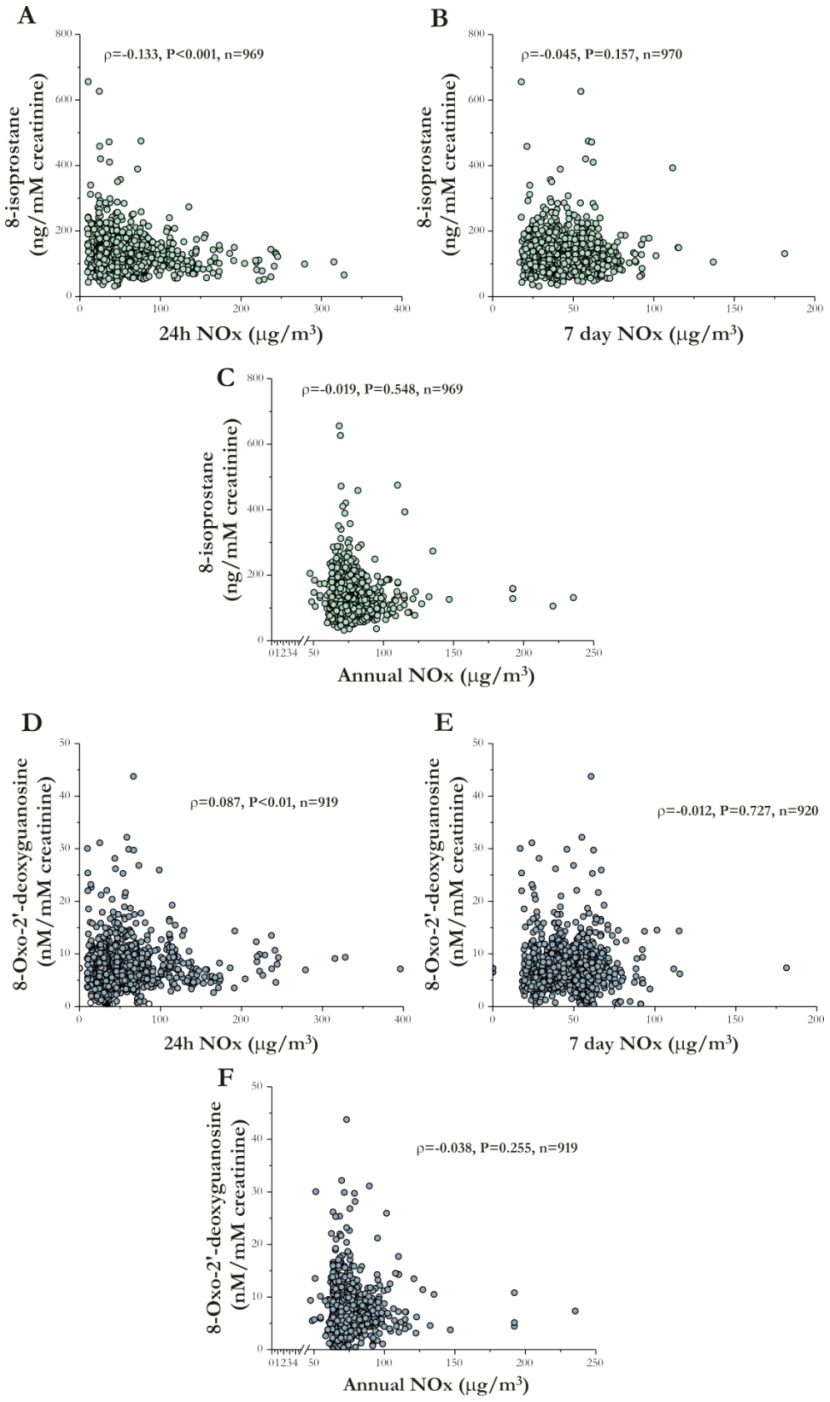


Figure 5.6 Correlations between urinary biomarkers and NOx exposures**.** The association between exposures and urinary biomarkers were examined using the Spearman Rank Order correlation. For each panel, the number of samples included in the analysis 'n', the correlation coefficient 'ρ' and the level of significance of the correlation 'P' is listed in the upper right hand corner of the image. **Panels A-C** illustrates the associations between 8-isoP and the various modeled exposures to NOx, panels D-F, the associations with 8-oxodG.

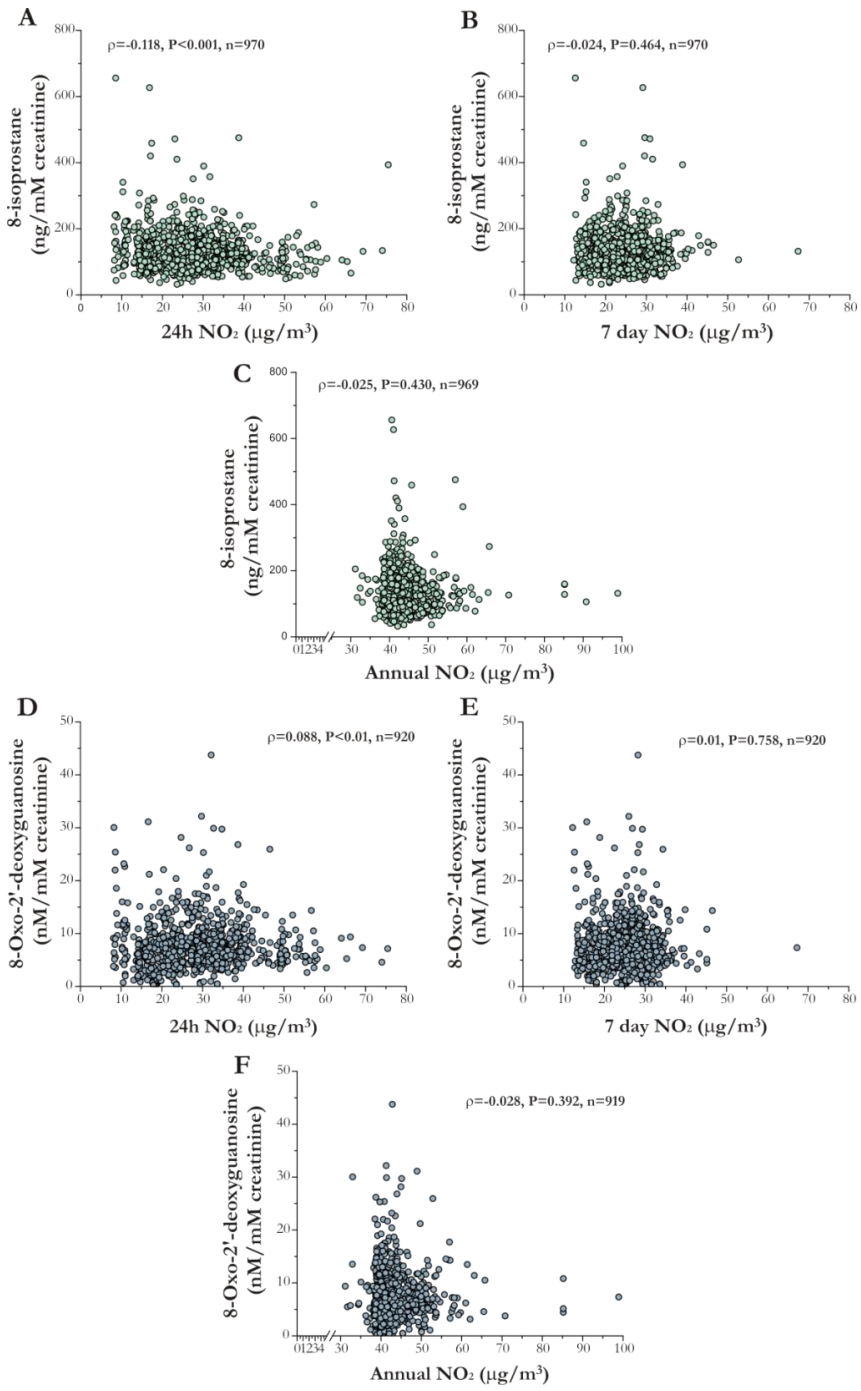


Figure 5.7 Correlations between urinary biomarkers and NO_2 exposures. The association between exposures and urinary biomarkers were examined using the Spearman Rank Order correlation. For each panel, the number of samples included in the analysis 'n', the correlation coefficient 'ρ' and the level of significance of the correlation 'P' is listed in the upper right hand corner of the image. **Panels A-C** illustrates the associations between 8-isoP and the various modeled exposures to NO_2, **Panels D-F**, the associations with 8-oxodG.

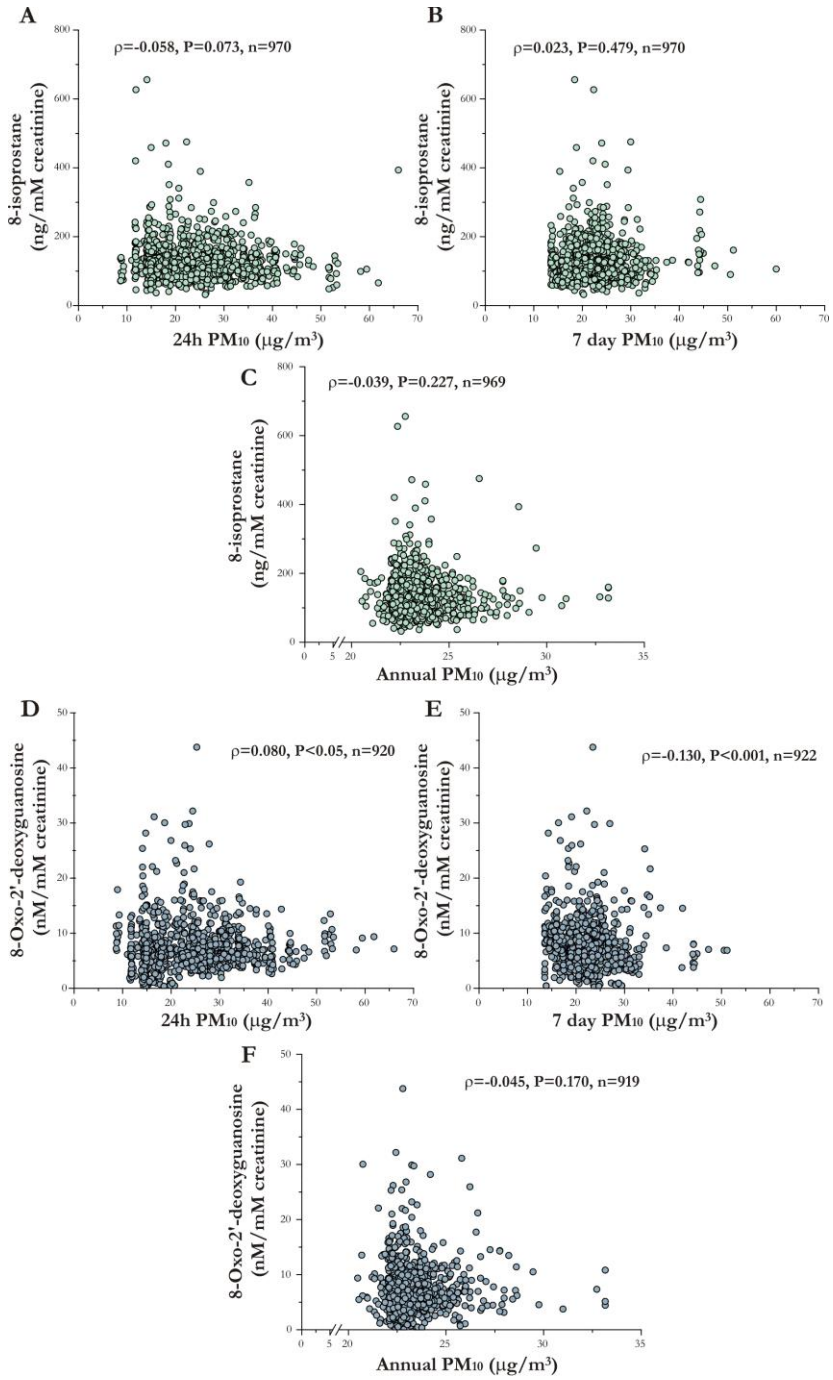


Figure 5.8 Correlations between urinary biomarkers and PM_{10} exposures. The association between exposures and urinary biomarkers were examined using the Spearman Rank Order correlation. For each panel, the number of samples included in the analysis 'n', the correlation coefficient 'ρ' and the level of significance of the correlation 'P' is listed in the upper right hand corner of the image. **Panels A-C** illustrates the associations between 8-isoP and the various modeled exposures to PM_{10}, **Panels D-F**, the associations with 8-oxodG.

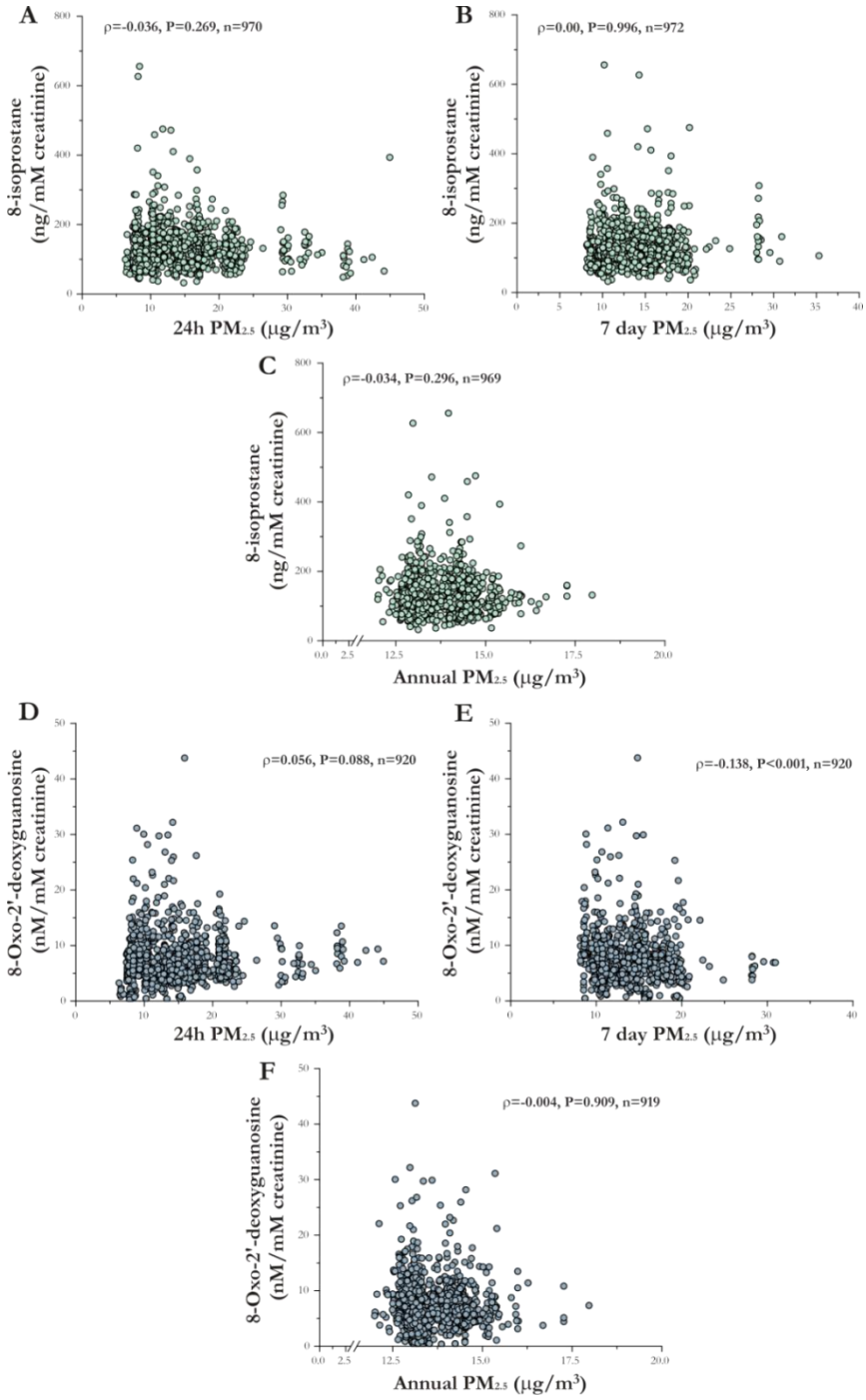


Figure 5.9 Correlations between urinary biomarkers and $PM_{2.5}$ exposures. The association between exposures and urinary biomarkers were examined using the Spearman Rank Order correlation. For each panel, the number of samples included in the analysis 'n', the correlation coefficient 'ρ' and the level of significance of the correlation 'P' is listed in the upper right hand corner of the image. **Panels A-C** illustrates the associations between 8-isoP and the various modeled exposures to $PM_{2.5}$, **Panels D-F**, the associations with 8-oxodG.

When these relationships were examined in the linear mixed modeled controlled for the covariants, previously outlined, significant negative associations were noted for 8-isoprostane with 24 hour, but not 7 day or annual modeled exposures for NOx and NO_2 equivalent to a 0.1 and 0.4% reduction in urinary 8-isoprostane per $\mu g/m^3$ increase in NOx and NO_2 respectively – **Table 5.4**. No associations were apparent for any of the modeled PM_{10} and $PM_{2.5}$ exposures with urinary 8-isoprostane concentrations. In contrast, urinary 8-oxodG concentrations were positively associated with the 24h modeled exposures for all four of the pollutants examined, with the associations for PM_{10} and $PM_{2.5}$ remaining significant after the correction for multiple testing. This equated to a 0.12% to 1.62% increased in urinary 8-oxodG per mg/m^3 PM_{10} and $PM_{2.5}$ respectively – **Table 5.4**. Conversely, there was a negative association between 8-oxodG and 7 day average NOx. No significant effects were apparent between the long term exposure to NOx, NO_2, PM_{10} and $PM_{2.5}$ estimates and 8-oxodG.

Table 5.4 Effect estimates (ratio of gemetric means) for the association between modeled air pollutant concentrations and oxidative stress biomarkers (8-isoprostane and 8-oxodG) concentrations

	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
8-isoprostane				
Annual	1.0005[0.9987,1.0022]	1.0013[0.9967,1.0060]	1.0025[0.9848,1.0205]	1.005[0.9626,1.0492]
7 days	0.999[0.9973,1.0007]	0.9992[0.9943,1.0041]	0.9965[0.9912,1.0019]	0.9947[0.9869,1.0025]
24 hour	₸0.9990**[0.9984,0.9997]	₸0.9963**[0.9937,0.9989]	0.9975[0.9946,1.0004]	0.9978[0.9935,1.0021]
8-oxodG				
Annual	0.9994[0.9963,1.0025]	0.9985[0.9902,1.0068]	0.9898[0.9599,1.0208]	0.9449[0.8774,1.0177]
7 days	₸0.9950**[0.9919,0.9981]	0.993[0.9843,1.0017]	0.9921[0.9822,1.0022]	0.9844*[0.9698,0.9993]
24 hour	1.0012*[1.0000,1.0024]	1.0047*[1.0001,1.0093]	₸1.0084**[1.0031,1.0137]	₸1.0162***[1.0079,1.0245]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.05, **p<0.01, ***p<0.001. ₸ Significant after the multiple testing procedures has been applied.

5.3.3 Descriptive Statistics of Urinary Metal Concentrations

Fifteen urinary metals were analyzed to investigate their predictive value as biomarkers of traffic exposure. **Figure 5.10** illustrates the overall data for the investigated metals in all subjects, in addition to stratification by exposure to ETS (environmental tobacco smoke). Antimony (Sb) and tungsten (W) showed values below detection limit, whereas titanium (Ti) showed values higher than Zn, which was considered unrealistic, suggesting a possible unknown isotopic interference at the 46.95, potentially as $^{14}N^{16}O_2^+$ or $^{12}C^{35}Cl^+$. Zinc (Zn) showed the highest urinary concentration, followed by a panel on medium abundance metals, Cu, Fe, Mo and As, with trace concentrations of Ba, Mn, Ni, Cr, Cd, V and Sn. The lower panel of **Figure 5.10** illustrates differences in metal concentration between ETS-exposed and non-ETS exposed subjects. There was a noticeable pattern of increased urinary metal concentrations associated with ETS exposure: for Ni, 9.9 (5.8 - 17.4) versus 10.9 (6.5 - 19.2) nM/mM creatinine, $p=0.06$; Ba 4.81 (2.86 - 8.79) versus 6.18 (3.74 - 9.54) nM/mM creatinine, $p<0.01$); Mn 2.79 (1.36 - 6.76) versus 6.18 (3.74 - 9.54) nM/mM creatinine, $p<0.05$) and Sn 2.56 (1.37 - 5.23) versus 3.65 (2.03 to 7.86) nM/mM creatinine, $p<0.001$. Molybdenum in contrast was significantly lower in the subjects with evidence of ETS exposure: 153.3 (102.9 - 210.9), $p<0.01$, versus 170.93 (118.65 - 246.08) nM/mM creatinine, $p<0.01$.

Crude correlations between the measured urinary metals and the modeled pollutant concentrations are illustrated in **Table 5.5.** The correlation matrix revealed a number of significant interactions between the various metals: 71 significant associations from 144 potential interactions. Overall, many of the observed significant associations were either weakly positive or negative, with only the trace elements of oil combustion (Cr and V) being strongly associated with chronic exposure to $PM_{2.5}$: For Cr vs. annual $PM_{2.5}$: $\rho=0.469$, $p<0.001$; and for V vs. annual $PM_{2.5}$: $\rho=0.440$, $p<0.001$.

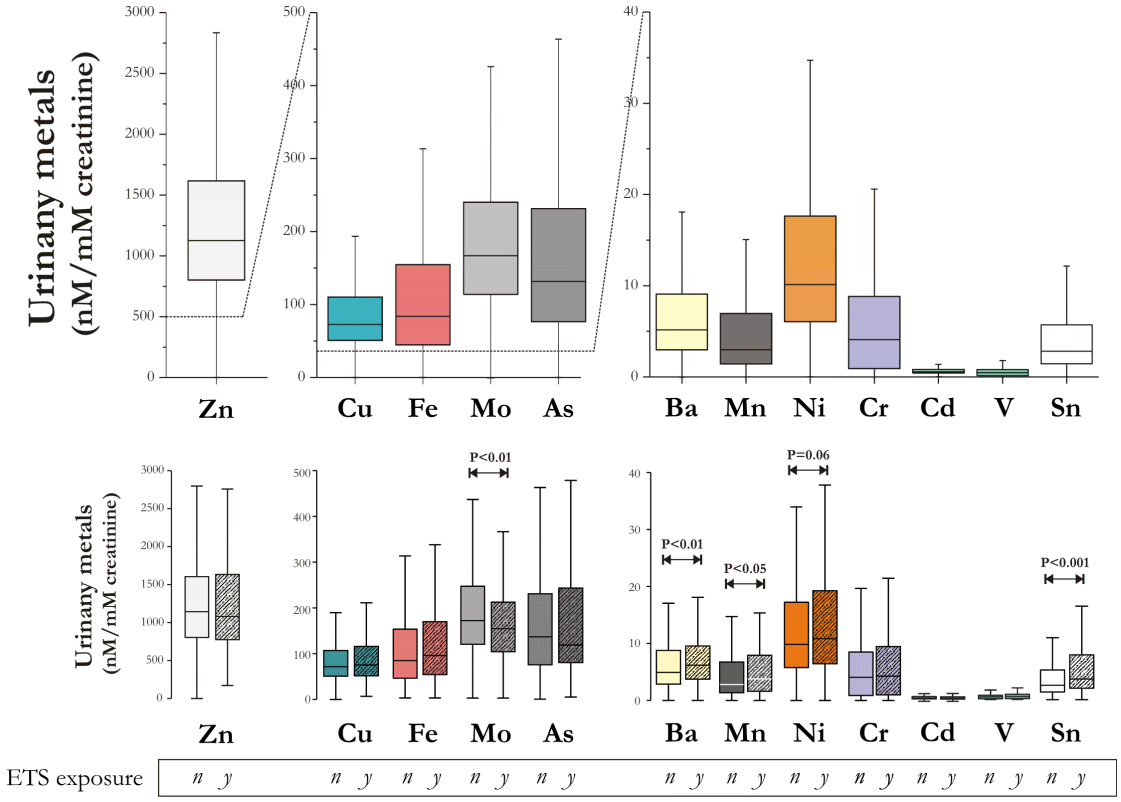


Figure 5.10 Urinary metal concentrations in school children. The data in the upper panel is associated with the overall study population stratified by concentration. The lower panel classifies the data by exposure to ETS. Data are presented as box plots, as described in the legend of **Figure 5.4.** Comparison between the ETS and non-ETS exposed groups was performed using the Mann-Whitney U test. Notes: *n*= no (not exposed to ETS), *y*=yes (exposed to ETS).

For the subsequent analyses, I focused only on Cu and Ni, as biomarkers of vehicular abrasion and oil combustion. **Figures 5.11** and **5.12** illustrates urinary concentrations of Cu (vehicular abrasion) and Ni (exhaust emission) respectively, based on school location, ethnicity and sex. Significant geographical differences ($p<0.001$) in the urinary concentrations of the two metals were observed, grouping the year 1-3 volunteers by school location (**Figures 5.11** and **5.12**, **Panel A**). Significant differences in urinary metal concentrations were observed between the different ethnic groups. The results largely concurred with oxidative stress biomarkers analysis, with children of Asian ethnicity having significantly higher urinary metal concentrations than Black and

White, **Figures 5.11** and **5.12**, **Panel C**. Female subjects also showed significantly higher levels of both metals compared to males (**Figures 5.11** and **5.12**, **Panel D**). The correlation between urinary Cu concentrations and the measured biomarkers of oxidative stress demonstrated weak through significant positive associations; for 8-isoprostane (ρ= 0.149, p<0.01, n= 953 sample) and 8-oxodG (ρ = 0.106, p=0.001, n=911) – **Figure 5.11**, **Panel B**. Similar, though weaker positive associations were also noted between Ni and 8-isoprostane and 8-oxodG (**Figure 5.12**, **Panel B**).

Table 5.5 The correlation matrix for the attribution of urinary metals to the modeled pollutants

	NOx			NO_2			PM_{10}			$PM_{2.5}$		
	24 h	7 day	Year	24 h	7 day	Year	24 h	7 day	Year	24 h	7 day	Year
Cu	-0.008	-0.092**	0.064	-0.014	-0.099**	0.051	0.122***	0.088**	0.094**	0.116***	0.054	0.165***
	951	952	951	952	952	951	952	952	951	952	954	951
Fe	0.095**	0.025	-0.052	-0.148***	-0.059	-0.064*	-0.102**	0.105**	-0.003	-0.113***	0.159***	0.088**
	951	952	951	952	952	951	952	952	951	952	954	951
Ba	-0.119***	-0.143***	0.086**	-0.110***	-0.135***	0.074*	0.014	-0.040	0.130***	0.013	-0.054	0.212***
	951	952	951	952	952	951	952	952	951	952	954	951
Mn	-0.061	-0.056	0.144***	-0.100***	-.096**	0.129**	-0.026	0.072*	0.221***	-0.048	0.091**	0.353***
	951	952	951	952	952	951	952	952	951	952	954	951
Mo	-0.057	-0.061	-0.087**	-0.059	-0.096**	-0.101**	-0.008	-0.007	-0.081*	-0.014	0.014	-0.041
	951	952	951	952	952	951	952	952	951	952	954	951
Zn	-0.034	-0.005	-0.186***	-0.046	-0.037	-0.191***	-0.009	0.063	-0.204***	0.005	0.075*	-0.212***
	951	952	951	952	952	951	952	952	951	952	954	951
Sn	-0.056	0.008	-0.063	-0.060	-0.033	0.067*	-0.023	0.020	-0.041	-0.011	0.057	0.004
	951	952	951	952	952	951	952	952	951	952	954	951
Cd	-0.027	-0.070*	0.010	-0.041	-0.111***	-0.005	0.049	-0.027	0.063	0.025	0.001	0.157***
	951	952	951	952	952	951	952	952	951	952	954	951
As	-0.051	-0.068*	0.030	-0.052	-0.068*	0.021	-0.025	-0.049	0.091**	-0.018	-0.022	0.199***
	951	952	951	952	952	951	952	952	951	952	954	951
Ni	-0.074*	0.019	0.007	-0.145***	-0.072*	0.003	-0.080*	0.039	0.070*	-0.078*	0.113***	0.158***
	951	952	951	952	952	951	952	952	951	952	954	951
Cr	-0.127***	-0.201***	0.184***	-0.164***	-0.227***	0.166***	-0.024	-0.038	0.290***	-0.055	-0.023	0.469***
	951	952	951	952	952	951	952	952	951	952	954	951
V	-0.056	-0.014	0.156***	-0.104**	-0.081*	0.141***	-0.038	0.040	0.254***	-0.044	0.088**	0.440***
	951	952	951	952	952	951	952	952	951	952	954	951

Notes: The upper number indicates the correlation coefficient between the metal with the exposure parameter; the lower number (blue) number of values included in the correlation. Significant correlations are highlighted: blue cells, *=p<0.05; orange cells, **=p<0.001; and green cells, ***= p<0.001.

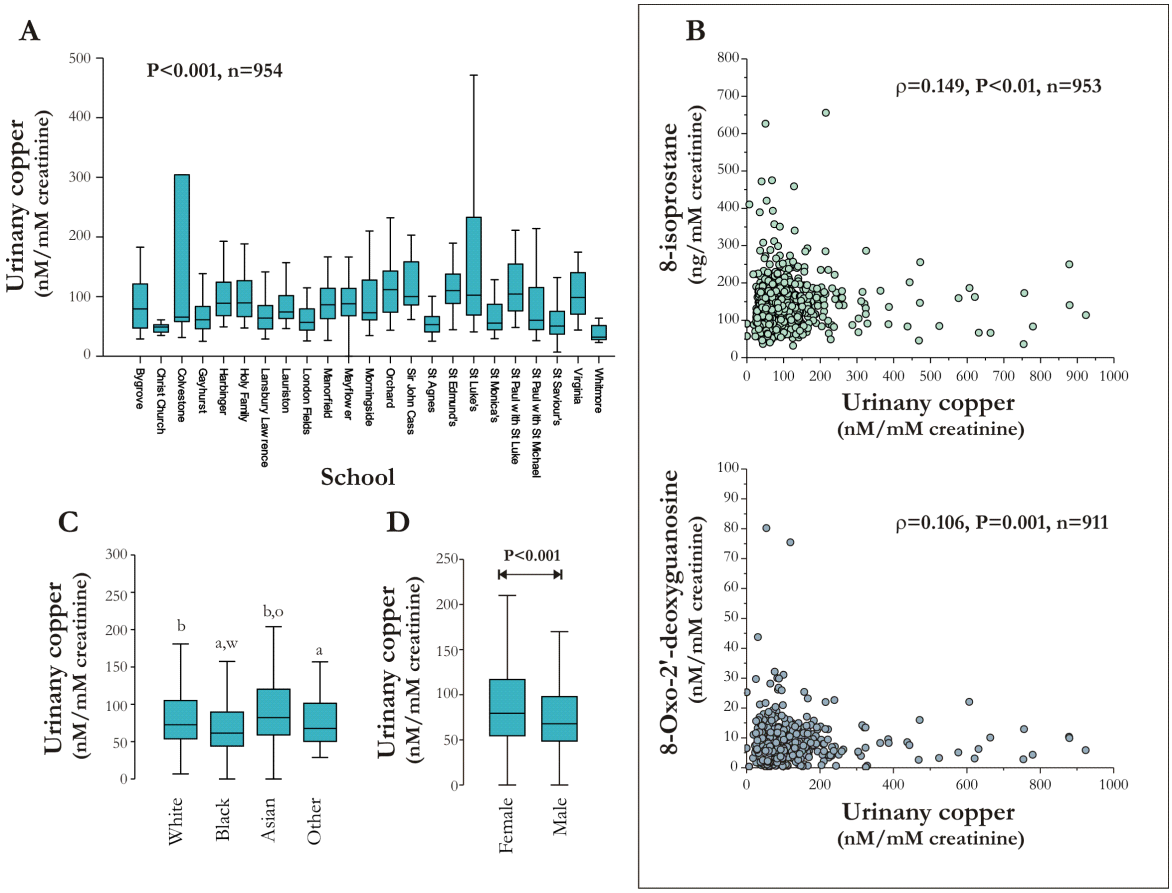


Figure 5.11 Descriptive analyses of urinary Copper. **Panel A** illustrates the geographical variation in urinary Cu concentration by school location. **Panel B** illustrated the association between urinary copper with 8-isoprostane and 8-Oxo-2'-deoxyguanosine concentrations. Difference in urinary Cu concentrations between different ethnicities and sex are illustrated in **Panels C** and **D** respectively. Details of analysis are identical to those described in the legend of **Figure 5.4.**

The adjusted associations between the measured urinary copper and nickel concentrations with modeled NOx, NO_2, PM_{10} and $PM_{2.5}$ over the 24h, 7 days and annual periods are presented in **Table 5.6**. Urinary Cu was not associated with the annual modeled exposures of any of the pollutants; however, significant negative associations were noted for 7 day average NOx (0.02% per $\mu g/m^3$) and 24 hour $PM_{2.5}$ (0.06%, per $\mu g/m^3$) exposures. In contrast, urinary nickel was significantly decreased by 0.3% and 0.7% when related to the 7 day average PM_{10} and $PM_{2.5}$ respectively. No

significant associations were noted between urinary Ni and long term pollution exposures.

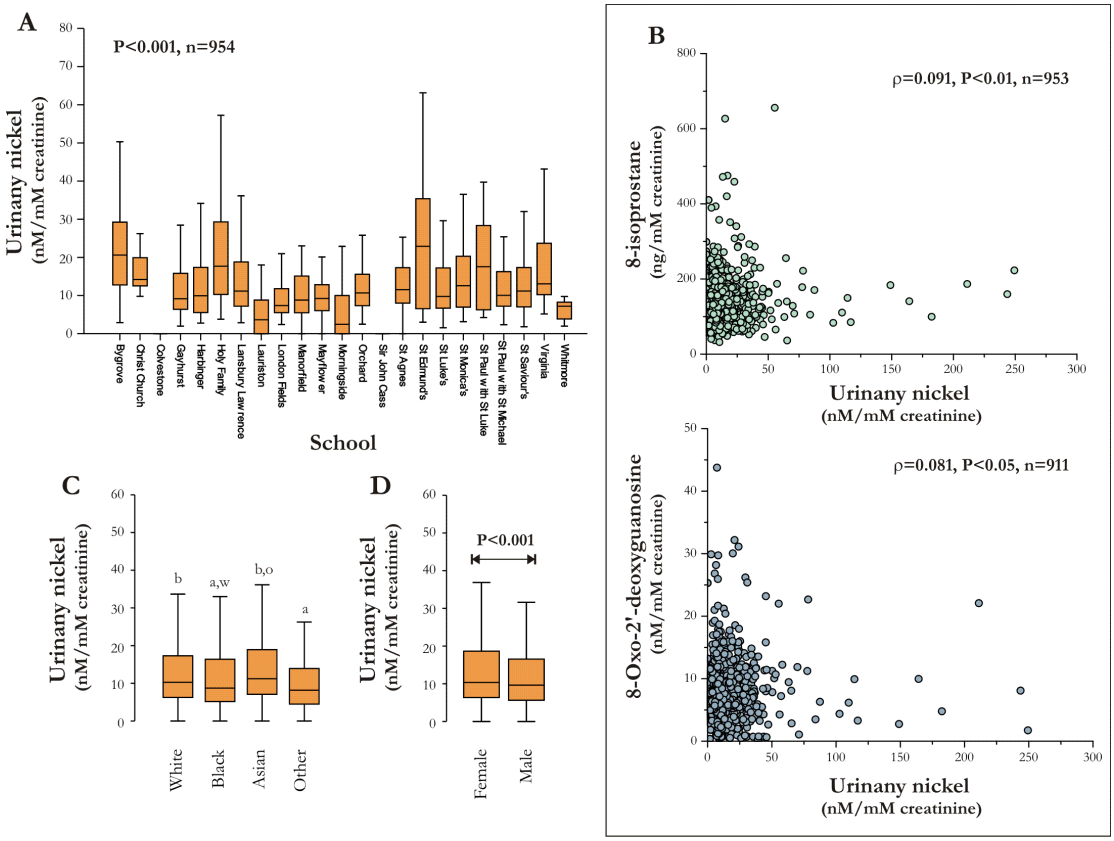


Figure 5.12 Descriptive analyses of urinary Nickel. **Panel A** illustrates the geographical variation in urinary Ni concentration by school location. **Panel B** illustrated the association between urinary Ni with, 8-isoprostane and 8-oxo-2'-deoxyguanosine concentrations. Differences in urinary Ni concentrations between ethnicities and sex are illustrated in **Panels C** and **D** respectively. Details of analysis are identical to those described in the legend of **Figure 5.4.**

5.3.4 Genetics Analysis

5.3.4.1 Allele and Genotype Frequencies

The distribution of the genotype and allele frequencies for Nrf2, GCLM and SOD3 SNPs for the total population and the population stratified by ethnicity is shown in **Tables J2** and **J3** (**Appendix J**). The genotype distributions between the ethnicities

for Nrf2 were only similar for one SNP, rs2001350 (A/G) with a minor allele frequency (MAF) of ~12%. For SNP rs2364723 (G/C), Black, White and 'Others' shared the same common genotype, 'GG', however, there was little variation amongst the ethnicities, with MAF of 20%, 28% and 36% each. In contrast, 'C' represented the major allele for the Asian population with MAF of 44%. Whereas for SNP rs7557529 (T/C), although Asian, Black and 'Others' shared the same common genotype, there was high heterozygosity for Black and 'Others'.

Table 5.6 Effect estimates (% change) for the association between urinary Cu and Ni concentrations with annual, 7 day and 24 hour modeled pollutant exposures

	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
Copper				
Annual	1.0000[0.9999,1.0002]	1.0000[0.9996,1.0004]		1.0019[0.9983,1.0055]
7 days	₸0.9998**[0.9997,0.9999]	0.9999[0.9996,1.0003]	0.9997[0.9993,1.0001]	0.9997[0.9991,1.0003]
24 hour	1.0000[0.9999,1.0000]	0.9999[0.9997,1.0001]	0.9998[0.9996,1.0001]	₸0.9994***[0.9991,0.9997]
Nickel				
Annual	1.0005[0.9998,1.0012]		1.0040[0.9971,1.0110]	1.0150[0.9982,1.0320]
7 days	0.9998[0.9992,1.0003]	1.0011[0.9996,1.0026]	₸0.9968**[0.9948,0.9988]	₸0.9930***[0.9902,0.9958]
24 hour	0.9999[0.9998,1.0001]	1.0009*[1.0001,1.0017]	1.0002[0.9993,1.0011]	0.9998[0.9984,1.0011]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.05, **p<0.01, ***p<0.001. ₸ Significant after the multiple testing procedures has been applied. Notes: Blank cells indicate non-converged data.

For GCLM SNP rs2301022 (G/A), all the ethnicities shared the same common genotype 'GG' with a MAF between 29% to 38%, except for the Black population, where 'AA' was the common genotype with a MAF of 47%. For SNP rs3170633 (G/A), all ethnicities shared the same common genotype 'GG' with Asians having ~15% higher MAF. In contrast, 'AA' was the major allele for Black with a MAF of 30%. In addition, all the ethnicities shared the same common genotype for all SOD3 SNPs, except rs699473 (C/T). While the Asian and 'Others' populations had a similar allele frequency of 'C', 50%, allele 'T' was the minor allele for Black population (25%) and 'C' is minor allele for White group (38%).

5.3.4.2 Assessing the Modifying Effect of Genotype on the Relationship between Biomarkers of Oxidative Stress and Ambient Pollutant Exposures

Tables 5.7 and **5.8** illustrate the crude and adjusted effects of Nrf2, GCLM and SOD3 genotypes on urinary 8-isoprostane and 8-oxodG concentrations. While unadjusted and adjusted estimates suggested significant associations with some of the genotypes, with 8-isoprostane (Nrf2, rs7557529) and 8-oxodG (Nrf2, rs7557529, rs2001350), these interactions were no longer significant after correction for multiple testing. This indicates that the studied SNPs in Nrf2, GCLM and SOD3 do not have direct effects on measured biomarkers of oxidative stress.

Table 5.7 Unadjusted and adjusted effect estimates (β) with 95% confidence intervals (95% CI) for the association between biomarkers of oxidative stress and Nrf2 and GCLM genotypes

Genotype	8-isoprostane		8-oxodG	
	Unadjusted (n=943)	Adjusted (n=932)	Unadjusted (n=894)	Adjusted (n=886)
		Nrf2		
rs2364723 G/C				
GG	Reference	Reference	Reference	Reference
GC	1.0167[0.9619,1.0747]	1.0076[0.9527,1.0658]	0.9923[0.9038,1.0895]	0.9767[0.8891,1.0731]
CC	1.0834*[1.0056,1.1674]	1.0305[0.9523,1.1152]	1.033[0.9098,1.1728]	1.0185[0.8922,1.1627]
GC+CC	1.0338[0.9815,1.0888]	1.0128[0.9600,1.0684]	1.0027[0.9185,1.0946]	0.986[0.9014,1.0785]
rs7557529 T/C				
TT	Reference	Reference	Reference	Reference
TC	0.9217**[0.8723,0.9738]	0.9405*[0.8899,0.9940]	0.9069*[0.8269,0.9946]	0.8997*[0.8205,0.9866]
CC	0.9374[0.8689,1.0113]	0.9476[0.8772,1.0236]	1.0514[0.9251,1.1950]	1.0329[0.9078,1.1752]
TC+CC	0.9255**[0.8787,0.9748]	0.9422*[0.8940,0.9930]	0.94[0.8612,1.0260]	0.929[0.8509,1.0143]
rs2001350 A/G				
AA	Reference	Reference	Reference	Reference
AG	0.9814[0.9227,1.0439]	0.9805[0.9227,1.0418]	0.9597[0.8656,1.0639]	0.9585[0.8660,1.0608]
GG	0.9278[0.7260,1.1858]	0.9497[0.7460,1.2090]	1.4541[0.9751,2.1683]	1.4921*[1.0073,2.2103]
AG+GG	0.9788[0.9213,1.0399]	0.979[0.9224,1.0391]	0.9797[0.8853,1.0841]	0.9795[0.8866,1.0822]
		GCLM		
rs2301022 G/A				
GG	Reference	Reference	Reference	Reference
GA	0.9508[0.9001,1.0043]	0.9647[0.9131,1.0191]	0.9367[0.8540,1.0274]	0.9484[0.8648,1.0401]
AA	0.9605[0.8897,1.0369]	0.9848[0.9105,1.0651]	0.8429**[0.7419,0.9575]	0.8938[0.7846,1.0182]
GA+AA	0.9532[0.9055,1.0034]	0.9692[0.9200,1.0210]	0.9118*[0.8360,0.9945]	0.9354[0.8570,1.0210]
rs3170633 G/A				
GG	Reference	Reference	Reference	Reference
GA	0.9876[0.9326,1.0458]	1.0038[0.9469,1.0641]	0.9359[0.8501,1.0303]	0.9602[0.8710,1.0585]
AA	0.9907[0.9216,1.0650]	1.0379[0.9578,1.1248]	0.8898[0.7857,1.0077]	0.988[0.8617,1.1328]
GA+AA	0.9885[0.9373,1.0424]	1.0111[0.9560,1.0695]	0.9228[0.8432,1.0099]	0.9658[0.8790,1.0611]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.5, **p<0.01.

Table 5.8 Unadjusted and adjusted effect estimate (β) and 95% confidence intervals (95% CI) for the association between biomarkers of oxidative stress and SOD3 genotypes

Genotype	8-isoprostane		8-oxodG	
	Unadjusted (n=941)	Adjusted (n=930)	Unadjusted (n=892)	Adjusted (n=884)
rs8192287 G/T				
GG	Reference	Reference	Reference	Reference
GT	1.0194[0.9502,1.0936]	1.0029[0.9338,1.0772]	1.0488[0.9316,1.1806]	1.0055[0.8922,1.1332]
TT	0.8661[0.6689,1.1214]	0.8889[0.6888,1.1472]	1.1656[0.7653,1.7753]	1.0473[0.6912,1.5868]
GT+TT	1.0098[0.9429,1.0815]	0.9958[0.9287,1.0677]	1.0555[0.9405,1.1846]	1.0081[0.8972,1.1327]
rs699473 C/T				
CC	Reference	Reference	Reference	Reference
CT	0.9829[0.9270,1.0421]	0.9582[0.9028,1.0171]	1.0194[0.9243,1.1244]	0.9895[0.8950,1.0941]
TT	1.0266[0.9580,1.1000]	1.0087[0.9380,1.0847]	1.1499*[1.0245,1.2907]	1.0721[0.9491,1.2109]
CT+TT	0.9972[0.9444,1.0531]	0.9727[0.9191,1.0293]	1.0614[0.9687,1.1630]	1.0131[0.9209,1.1146]
rs13306703 C/T				
CC	Reference	Reference	Reference	Reference
CT	0.9763[0.9229,1.0329]	0.9851[0.9307,1.0426]	0.9953[0.9055,1.0941]	1.0079[0.9168,1.1081]
TT	1.0236[0.9037,1.1594]	1.0359[0.9149,1.1729]	1.0524[0.8585,1.2901]	1.0723[0.8754,1.3135]
CT+TT	0.9824[0.9310,1.0366]	0.9917[0.9395,1.0468]	1.003[0.9166,1.0974]	1.0166[0.9290,1.1125]
rs8192288 G/T				
GG	Reference	Reference	Reference	Reference
GT	1.027[0.9579,1.1010]	1.0106[0.9415,1.0848]	1.0487[0.9324,1.1795]	1.0043[0.8918,1.1310]
TT	0.8665[0.6695,1.1214]	0.8898[0.6898,1.1478]	1.1652[0.7653,1.7739]	1.0451[0.6900,1.5831]
GT+TT	1.017[0.9502,1.0886]	1.0031[0.9361,1.0749]	1.0554[0.9411,1.1835]	1.0067[0.8966,1.1304]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.5, **p<0.01.

Subsequently, the interaction effects of Nrf2, GCLM and SOD3 SNPs on the relationship between pollutant exposure and oxidative stress biomarkers were examined. After adjustment for multiple comparisons, no significant interactions of the studied polymorphisms were observed on the relationship between urinary oxidation markers and the acute 24 hour exposure attributions – **Appendix J, Tables J4**– **J7**. In contrast, based on the 7 day (sub-chronic) exposure attribution, SNPs in Nrf2 were found to have positive effect on the urinary concentrations of 8-isoP (rs7557529) in relation to NO_2 concentrations, and 8-oxodG (rs2364723), with respect to PM_{10} – **Appendix J, Tables J8** and **J9** respectively. In the former case (rs7557529 – Nrf2) possession of the 'C' allele was associated with significant 1.5% increase in urinary 8-IsoP concentrations per $\mu g/m^3$ NO_2 in the heterozygote's compared with the reference 'TT' genotype (**Figure 5.13**). Similar associations were observed with the 7-day NOx exposure attributions, but these failed to attain statistical significance after adjustment

for multiple testing (**Figure 5.13**). When these data were examined stratified according to the three major ethnic groups it was evident that the positive effect of the 'C' allele was largely restricted to the black children for the 7 day exposures to NOx and NO_2, and only seen in the heterozygotes (TC) and in the combined model (TC+CC) (**Table 5.8**). In addition, these was some evidence that the 'C' allele was associated with lower urinary 8-isoP concentrations in relation to annual exposures to PM_{10} and $PM_{2.5}$, which was not apparent in the combined analysis.

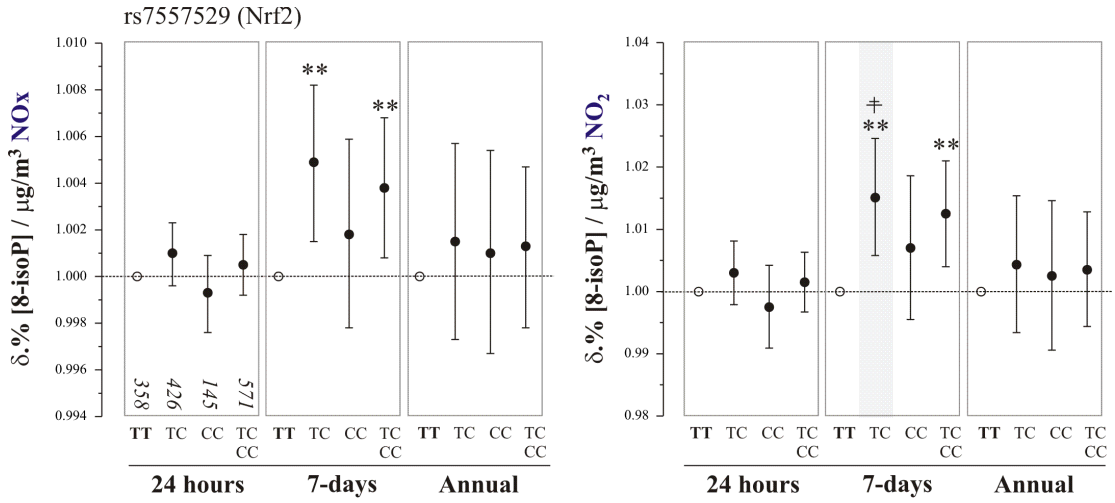


Figure 5.13 Influence of rs7557529 (Nrf2) on the association between urinary 8-isoP concentrations and 24 hour, 7-day or annual pollutant attributions to NOx and NO_2. Data are expressed relative to the major allele across the whole multi-ethnic population-TT (**Appendix J**, **Table J2**). Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year, including effect for school location. Nominally significant differences relative to the reference genotype are illustrated as follows: #p<0.1, *p<0.05 and **p<0.01. Where these differences remained significant (P<0.05) after adjusted for multiple comparisons this is illustrated by 'ǂ'.

For rs2364723 (Nrf2) possession of the 'C' allele, in the combined model was associated with a 2.8% increase in urinary 8-oxodG concentrations per unit increase in the sub-chronic (7 day) modeled PM_{10} concentrations compared with the GG reference genotype (**Figure 5.14**). A similar pattern was observed with 7 day exposure attributions to both NO_2 and $PM_{2.5}$, but these failed to reach statistical significance after

correction for multiple comparisons. Similar, though non-significant trends were also observed with this pollutant using the acute 24h exposure attributions (**Figure 5.14**).

Table 5.8 Influence of rs7557529 (Nrf2) on the association between urinary 8-isoP concentrations and 24 hour, 7-day or annual pollutant attributions to NOx and NO_2 for each of the major ethnicities within the study population. Data are expressed relative to the major allele for each population.

Genotype	Subjects	NOx 24 hours		NO_2 24 hours		PM_{10} 24 hours		$PM_{2.5}$ 24 hours	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (T/C)									
TT	176	*Reference*		*Reference*		*Reference*		*Reference*	
TC	128	1.0005	[0.9977,1.0033]	1.0004	[0.9914,1.0095]	1.0005	[0.9906,1.0105]	0.9987	[0.9836,1.0141]
CC	26	0.9983	[0.9955,1.0013]	**0.9853***	**[0.9722,0.9985]**	0.9883	[0.9726,1.0043]	0.9777	[0.9551,1.0007]
TC+CC	154	0.9995	[0.9972,1.0018]	0.9964	[0.9882,1.0047]	0.9978	[0.9887,1.0071]	0.9937	[0.9797,1.0078]
Black (T/C)									
TT	79	*Reference*		*Reference*		*Reference*		*Reference*	
TC	116	1.0007	[0.9977,1.0037]	1.0027	[0.9914,1.0141]	1.0063	[0.9935,1.0192]	1.008	[0.9898,1.0266]
CC	39	1.0032	[0.9986,1.0079]	1.0128	[0.9968,1.0291]	1.0086	[0.9905,1.0271]	1.0118	[0.9864,1.0379]
TC+CC	155	1.0011	[0.9982,1.0040]	1.0049	[0.9940,1.0159]	1.0069	[0.9946,1.0193]	1.0092	[0.9918,1.0268]
White (C/T)									
CC	68	*Reference*		*Reference*		*Reference*		*Reference*	
CT	117	1.0017	[0.9997,1.0037]	1.0051	[0.9961,1.0143]	1.0032	[0.9931,1.0135]	1.0054	[0.9914,1.0196]
TT	65	1.0013	[0.9987,1.0039]	1.0018	[0.9914,1.0124]	1.0011	[0.9886,1.0137]	1.0032	[0.9851,1.0216]
CT+TT	182	1.002	[0.9996,1.0035]	1.0038	[0.9954,1.0123]	1.0022	[0.9926,1.0119]	1.0043	[0.9907,1.0180]

Genotype	Subjects	NOx 7 days		NO_2 7 days		PM_{10} 7 days		$PM_{2.5}$ 7 days	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (T/C)									
TT	176	*Reference*		*Reference*		*Reference*		*Reference*	
TC	128	1.0022	[0.9969,1.0074]	1.0042	[0.9887,1.0199]	1.0087	[0.9953,1.0223]	1.0146	[0.9940,1.0357]
CC	26	0.995	[0.9873,1.0028]	0.981	[0.9558,1.0069]	0.9915	[0.9702,1.0133]	0.9827	[0.9507,1.0157]
TC+CC	154	1.0003	[0.9955,1.0051]	0.9992	[0.9849,1.0137]	1.0048	[0.9925,1.0172]	1.0071	[0.9883,1.0263]
Black (T/C)									
TT	79	*Reference*		*Reference*		*Reference*		*Reference*	
TC	116	1.0054	[0.9988,1.0121]	**1.0196***	**[1.0026,1.0368]**	0.9971	[0.9750,1.0198]	0.9926	[0.9602,1.0262]
CC	39	**1.0108***	**[1.0006,1.0211]**	**1.0288***	**[1.0034,1.0550]**	0.9962	[0.9624,1.0311]	0.9931	[0.9463,1.0422]
TC+CC	155	**1.0067***	**[1.0004,1.0130]**	**1.0218****	**[1.0056,1.0382]**	0.9977	[0.9768,1.0190]	0.9943	[0.9638,1.0258]
White (C/T)									
CC	68	*Reference*		*Reference*		*Reference*		*Reference*	
CT	117	1.0056	[0.9997,1.0114]	1.0165	[0.9994,1.0338]	0.9986	[0.9742,1.0236]	0.9987	[0.9640,1.0347]
TT	65	1.0005	[0.9929,1.0082]	0.9989	[0.9778,1.0204]	0.9781	[0.9488,1.0083]	0.9734	[0.9324,1.0163]
CT+TT	182	1.0041	[0.9987,1.0094]	1.011	[0.9956,1.0266]	0.9915	[0.9687,1.0148]	0.9899	[0.9574,1.0235]

Genotype	Subjects	NOx annual		NO_2 24 annual		PM_{10} annual		$PM_{2.5}$ annual	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (T/C)									
TT	176	*Reference*		*Reference*		*Reference*		*Reference*	
TC	128	1.0027	[0.9946,1.0109]	1.0067	[0.9861,1.0278]	1.0283	[0.9529,1.1096]	1.0325	[0.9162,1.1636]
CC	26	0.9898	[0.9724,1.0075]	0.9742	[0.9316,1.0187]	0.9637	[0.8187,1.1344]	0.952	[0.7669,1.1817]
TC+CC	154	1.0007	[0.9932,1.0083]	1.0015	[0.9823,1.0210]	1.018	[0.9480,1.0933]	1.0173	[0.9093,1.1381]
Black (T/C)									
TT	79	*Reference*		*Reference*		*Reference*		*Reference*	
TC	116	0.9933	[0.9849,1.0018]	0.9839	[0.9624,1.0059]	**0.9256***	**[0.8584,0.9980]**	**0.8582***	**[0.7539,0.9769]**
CC	39	1.0072	[0.9931,1.0215]	1.0161	[0.9806,1.0529]	1.0489	[0.9263,1.1876]	0.9815	[0.8059,1.1954]
TC+CC	155	0.9958	[0.9877,1.0039]	0.9898	[0.9692,1.0109]	0.9464	[0.8809,1.0167]	**0.8791***	**[0.7766,0.9952]**
White (C/T)									
CC	68	*Reference*		*Reference*		*Reference*		*Reference*	
CT	117	1.0032	[0.9962,1.0103]	1.0082	[0.9899,1.0268]	1.0177	[0.9508,1.0893]	0.9866	[0.8748,1.1127]
TT	65	1.0014	[0.9938,1.0090]	1.0021	[0.9823,1.0222]	1.0129	[0.9359,1.0962]	1.0346	[0.8995,1.1901]
CT+TT	182	1.0026	[0.9968,1.0084]	1.006	[0.9907,1.0216]	1.0164	[0.9559,1.0808]	1.0035	[0.8986,1.1207]

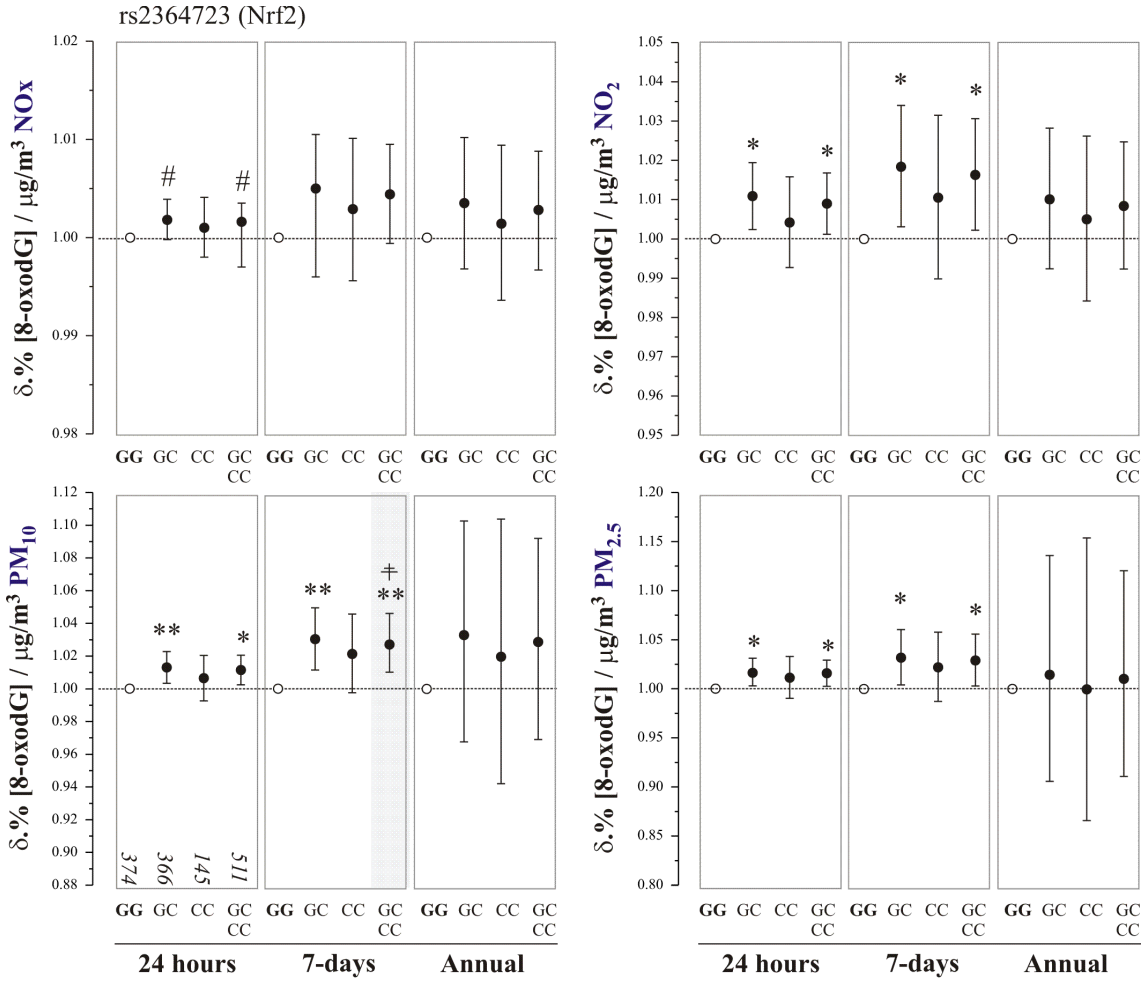


Figure 5.14 Influence of rs2364723 (Nrf2) on the association between urinary 8-oxodG concentrations and 24 hour, 7-day or annual pollutant attributions. Data are expressed relative to the major allele across the whole multi-ethnic population- GG (**Appendix J**, **Table J2**). Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year, including effect for school location. Nominally significant differences relative to the reference genotype are illustrated as follows: #p<0.1, *p<0.05 and ** p<0.01. Where these differences remained significant (P<0.05) after adjusted for multiple comparisons this is illustrated by '‡'.

When these analyses were repeated separately referenced againt the major allele for each ethnicity: Asian (C, 56%), black (G, 80%) and white (G, 72%), the 'C' allele was only associated with increased excretion of 8-oxodG in the white population (GC and GC+CC) for acute and sub-chronic exposures to PM_{10}, **Table 5.9**.

Table 5.9 Influence of rs2364723 (Nrf2) on the association between urinary urinary 8-oxodG concentrations and 24 hour, 7-day or annual pollutant attributions to NOx and NO_2 for each of the major ethnicities within the study population. Data are expressed relative to the major allele for each population

Genotype	Subjects	NOx 24 hours		NO_2 24 hours		PM_{10} 24 hours		$PM_{2.5}$ 24 hours	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (C/G)									
CC	105	*Reference*		*Reference*		*Reference*		*Reference*	
CG	145	1.0022	[0.9969,1.0074]	1.0093	[0.9911,1.0277]	1.0072	[0.9868,1.0279]	1.0092	[0.9789,1.0404]
GG	72	0.9972	[0.9918,1.0025]	0.9889	[0.9684,1.0098]	0.9845	[0.9606,1.0091]	0.9728	[0.9353,1.0118]
CG+GG	217	0.9997	[0.9952,1.0043]	1.0016	[0.9849,1.0184]	0.9994	[0.9805,1.0187]	0.9979	[0.9697,1.0269]
Black (G/C)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GC	79	1.0009	[0.9970,1.0050]	1.0046	[0.9891,1.0202]	1.002	[0.9845,1.0199]	0.9994	[0.9736,1.0260]
CC	8	0.9947	[0.9803,1.0094]	0.9648	[0.9116,1.0210]	0.9925	[0.9449,1.0426]	0.9967	[0.9221,1.0774]
GC+CC	87	1.0007	[0.9968,1.0046]	1.0029	[0.9878,1.0182]	1.0016	[0.9847,1.0189]	0.9998	[0.9747,1.0255]
White (G/C)									
GG	126	*Reference*		*Reference*		*Reference*		*Reference*	
GC	92	1.0013	[0.9985,1.0041]	1.0095	[0.9964,1.0228]	**1.0150***	**[1.0009,1.0294]**	1.0186	[0.9989,1.0387]
CC	19	1.0008	[0.9938,1.0079]	1.0037	[0.9799,1.0280]	1.0105	[0.9769,1.0451]	1.0045	[0.9497,1.0623]
GC+CC	111	1.001	[0.9985,1.0040]	1.0086	[0.9961,1.0212]	**1.0144***	**[1.0007,1.0284]**	1.0173	[0.9980,1.0370]

Genotype	Subjects	NOx 7 days		NO_2 7 days		PM_{10} 7 days		$PM_{2.5}$ 7 days	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (C/G)									
CC	105	*Reference*		*Reference*		*Reference*		*Reference*	
CG	145	0.9992	[0.9886,1.0098]	1.0003	[0.9693,1.0323]	1.0055	[0.9780,1.0339]	1.0046	[0.9635,1.0475]
GG	72	0.9878	[0.9751,1.0006]	0.9682	[0.9304,1.0075]	0.9682	[0.9318,1.0059]	0.9533	[0.9008,1.0088]
CG+GG	217	0.9956	[0.9857,1.0055]	0.9908	[0.9618,1.0207]	0.9968	[0.9705,1.0237]	0.9915	[0.9529,1.0317]
Black (G/C)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GC	79	1.0028	[0.9935,1.0123]	1.0105	[0.9875,1.0340]	1.0057	[0.9741,1.0383]	1.0001	[0.9537,1.0488]
CC	8	0.9819	[0.9586,1.0057]	0.9537	[0.8962,1.0148]	1.0245	[0.9376,1.1194]	1.0345	[0.9132,1.1719]
GC+CC	87	1.0009	[0.9920,1.0099]	1.006	[0.9838,1.0288]	1.0085	[0.9779,1.0401]	1.0054	[0.9608,1.0522]
White (G/C)									
GG	126	*Reference*		*Reference*		*Reference*		*Reference*	
GC	92	1.0044	[0.9951,1.0139]	1.0139	[0.9872,1.0413]	**1.0375***	**[1.0000,1.0763]**	1.0415	[0.9887,1.0971]
CC	19	1.0017	[0.9821,1.0217]	0.9853	[0.9317,1.0419]	1.0444	[0.9631,1.1327]	1.0392	[0.9387,1.1505]
GC+CC	111	1.004	[0.9955,1.0134]	1.0112	[0.9858,1.0372]	**1.0385***	**[1.0023,1.0759]**	1.0413	[0.9905,1.0946]

Genotype	Subjects	NOx annual		NO_2 24 annual		PM_{10} annual		$PM_{2.5}$ annual	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (C/G)									
CC	105	*Reference*		*Reference*		*Reference*		*Reference*	
CG	145	1.0046	[0.9937,1.0156]	1.0124	[0.9834,1.0422]	1.0575	[0.9444,1.1842]	1.1512	[0.9279,1.4283]
GG	72	1.0031	[0.9815,1.0252]	1.007	[0.9531,1.0639]	1.0557	[0.8741,1.2751]	1.2178	[0.9157,1.6195]
CG+GG	217	1.0044	[0.9938,1.0151]	1.0116	[0.9834,1.0406]	1.0572	[0.9479,1.1790]	1.1698	[0.9556,1.4320]
Black (G/C)									
GG	138	*Reference*		*Reference*		*Reference*		*Reference*	
GC	79	0.9969	[0.9851,1.0088]	0.9941	[0.9641,1.0251]	0.9649	[0.8670,1.0738]	0.9512	[0.7962,1.1364]
CC	8	0.9865	[0.9612,1.0126]	0.9647	[0.8998,1.0342]	0.91	[0.7148,1.1586]	0.8337	[0.5438,1.2782]
GC+CC	87	0.9958	[0.9845,1.0073]	0.9911	[0.9622,1.0209]	0.9596	[0.8657,1.0636]	0.9399	[0.7908,1.1170]
White (G/C)									
GG	126	*Reference*		*Reference*		*Reference*		*Reference*	
GC	92	1.0008	[0.9884,1.0134]	1.0015	[0.9699,1.0341]	0.9758	[0.8691,1.0957]	0.9224	[0.7585,1.1218]
CC	19	1.0085	[0.9938,1.0235]	1.0215	[0.9829,1.0615]	1.086	[0.9390,1.2560]	1.2069	[0.9026,1.6139]
GC+CC	111	1.004	[0.9939,1.0145]	1.0097	[0.9831,1.0370]	1.0149	[0.9177,1.1224]	0.9924	[0.8311,1.1850]

Of the SNPs examined in GCLM only SNP rs3170633 demonstrated a significant negative impact on the association between urinary 8-isoprostane concentrations and the 7 days exposures to $PM_{2.5}$ (**Appendix J, Table J8**). No interactions were observed between SNPs in GCLM and the relationship between air pollutant exposures and urinary 8-oxodG concentrations (**Appendix J, Table J8**). For rs3170633 children carrying the 'A' allele demonstrated a 2.2% reduction in urinary 8-isoprostane concentrations per unit $PM_{2.5}$. Interestingly, a similar, though non-significant negative association was noted when 24h exposures were considered (**Figure 5.15**). Examining associations by individual ethnic group in relation to their major allele: Asian (G, 81%), black (A, 70%) and white (64%), demonstrated that the AG genotype was associated with reduced excretion of 8-isoP in the white population, in relation to 24h exposures to PM_{10} and $PM_{2.5}$, **Table 5.10**. This effect was not significant for the homozygotes, or in the combined model, but in all cases the effect appeared protective, even though it failed to attain statistical significance.

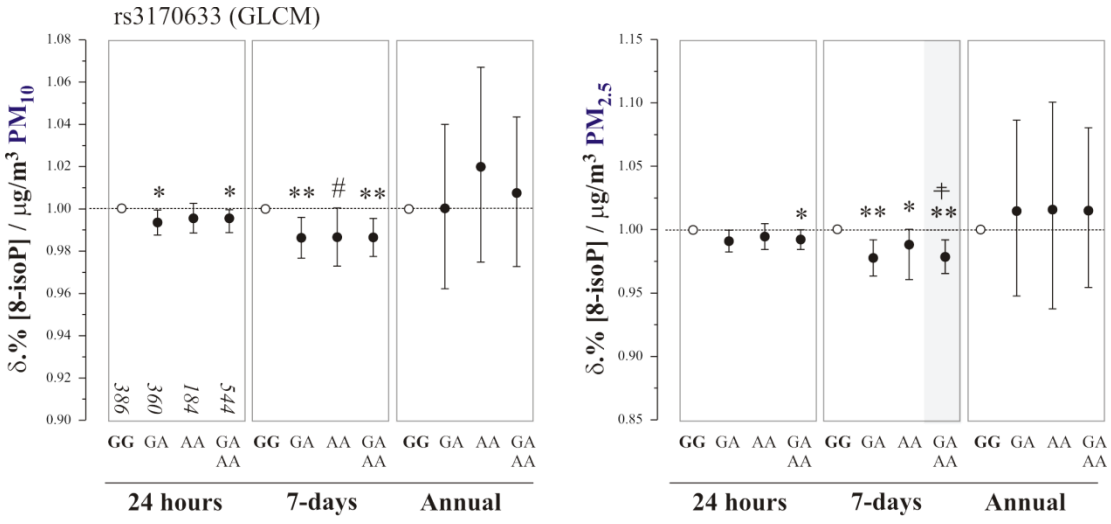


Figure 5.15 Influence of rs3170633 (GLCM) on the association between urinary 8-isoP concentrations and 24 hour, 7-day or annual pollutant attributions to PM_{10} and $PM_{2.5}$. Data are expressed relative to the major allele across the whole multi-ethnic population-GG (**Appendix J**, **Table J2**). Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year, including effect for school location. Nominally significant differences relative to the reference genotype are illustrated as follows: #p<0.1, *p<0.05 and **p<0.01. Where these differences remained significant (P<0.05) after adjusted for multiple comparisons this is illustrated by '‡'.

Table 5.10 Influence of rs3170633 (GLCM) on the association between urinary 8-isoP concentrations and 24 hour, 7-day or annual pollutant attributions to NOx and NO_2 for each of the major ethnicities within the study population. Data are expressed relative to the major allele for each population.

Genotype	Subjects	NOx 24 hours		NO_2 24 hours		PM_{10} 24 hours		$PM_{2.5}$ 24 hours	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/A)									
GG	216	*Reference*		*Reference*		*Reference*		*Reference*	
GA	105	1.0008	[0.9979,1.0036]	1.0029	[0.9936,1.0123]	1.0061	[0.9960,1.0163]	1.0092	[0.9942,1.0244]
AA	9	0.9928	[0.9800,1.0059]	0.9887	[0.9592,1.0192]	0.9929	[0.9606,1.0263]	0.9968	[0.9384,1.0588]
GA+AA	114	1.0004	[0.9976,1.0032]	1.0018	[0.9928,1.0109]	1.005	[0.9951,1.0149]	1.0082	[0.9935,1.0232]
Black (A/G)									
AA	119	*Reference*		*Reference*		*Reference*		*Reference*	
AG	92	0.9995	[0.9970,1.0020]	0.9968	[0.9869,1.0068]	1.0008	[0.9886,1.0131]	0.9997	[0.9816,1.0182]
GG	25	1.0001	[0.9947,1.0055]	1.0004	[0.9832,1.0179]	1.0137	[0.9932,1.0347]	1.0173	[0.9906,1.0448]
AG+GG	117	0.9996	[0.9971,1.0021]	0.9976	[0.9881,1.0071]	1.0036	[0.9921,1.0152]	1.0049	[0.9882,1.0219]
White (G/A)									
GG	99	*Reference*		*Reference*		*Reference*		*Reference*	
GA	116	0.9987	[0.9966,1.0008]	0.9933	[0.9847,1.0020]	**0.9898***	**[0.9802,0.9995]**	**0.9858***	**[0.9726,0.9991]**
AA	34	1.0007	[0.9985,1.0030]	1.0035	[0.9931,1.0141]	0.9981	[0.9866,1.0097]	0.9962	[0.9803,1.0123]
GA+AA	150	1.000	[0.9978,1.0015]	0.997	[0.9890,1.0050]	0.9928	[0.9839,1.0017]	0.9895	[0.9772,1.0019]

Genotype	Subjects	NOx 7 days		NO_2 7 days		PM_{10} 7 days		$PM_{2.5}$ 7 days	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/A)									
GG	216	*Reference*		*Reference*		*Reference*		*Reference*	
GA	105	0.9980	[0.9931,1.0030]	0.9939	[0.9792,1.0087]	**0.9851***	**[0.9734,0.9969]**	**0.9747****	**[0.9570,0.9928]**
AA	9	0.9881	[0.9691,1.0075]	0.9807	[0.9269,1.0376]	0.9722	[0.9143,1.0337]	0.9557	[0.8850,1.0320]
GA+AA	114	0.9976	[0.9927,1.0025]	0.9934	[0.9791,1.0080]	**0.9848***	**[0.9732,0.9965]**	**0.9740****	**[0.9564,0.9919]**
Black (A/G)									
AA	119	*Reference*		*Reference*		*Reference*		*Reference*	
AG	92	1.0007	[0.9941,1.0074]	1.0011	[0.9843,1.0183]	1.0073	[0.9851,1.0301]	1.0039	[0.9712,1.0378]
GG	25	1.002	[0.9903,1.0139]	1.004	[0.9750,1.0338]	**1.0557****	**[1.0145,1.0986]**	**1.0943****	**[1.0334,1.1587]**
AG+GG	117	1.0012	[0.9947,1.0076]	1.0025	[0.9864,1.0189]	1.0163	[0.9948,1.0382]	1.0222	[0.9903,1.0550]
White (G/A)									
GG	99	*Reference*		*Reference*		*Reference*		*Reference*	
GA	116	0.9972	[0.9905,1.0040]	0.9901	[0.9721,1.0085]	0.9935	[0.9700,1.0176]	0.9946	[0.9622,1.0281]
AA	34	1.0006	[0.9939,1.0075]	1.0039	[0.9836,1.0246]	1.0027	[0.9667,1.0401]	0.9991	[0.9487,1.0522]
GA+AA	150	0.999	[0.9932,1.0052]	0.9961	[0.9794,1.0131]	0.9961	[0.9735,1.0193]	0.9969	[0.9659,1.0289]

Genotype	Subjects	NOx annual		NO_2 24 annual		PM_{10} annual		$PM_{2.5}$ annual	
	(n)	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*	*b*	*ci95*
Asian (G/A)									
GG	216	*Reference*		*Reference*		*Reference*		*Reference*	
GA	105	1.0009	[0.9960,1.0058]	1.0026	[0.9891,1.0162]	1.0191	[0.9620,1.0795]	1.0818	[0.9688,1.2080]
AA	9	0.9964	[0.9830,1.0099]	0.9905	[0.9559,1.0263]	0.9767	[0.8513,1.1206]	1.0326	[0.8030,1.3277]
GA+AA	114	1.0006	[0.9958,1.0054]	1.0018	[0.9886,1.0151]	1.0153	[0.9604,1.0734]	1.0774	[0.9694,1.1975]
Black (A/G)									
AA	119	*Reference*		*Reference*		*Reference*		*Reference*	
AG	92	0.9954	[0.9865,1.0044]	0.9896	[0.9664,1.0133]	0.9628	[0.8919,1.0395]	0.9617	[0.8442,1.0956]
GG	25	0.9949	[0.9808,1.0092]	0.9878	[0.9522,1.0247]	0.9544	[0.8333,1.0931]	0.9292	[0.7367,1.1721]
AG+GG	117	0.9952	[0.9870,1.0035]	0.989	[0.9676,1.0108]	0.9607	[0.8929,1.0335]	0.9558	[0.8442,1.0822]
White (G/A)									
GG	99	*Reference*		*Reference*		*Reference*		*Reference*	
GA	116	0.9946	[0.9869,1.0023]	0.9867	[0.9672,1.0066]	0.9676	[0.9036,1.0362]	0.9626	[0.8559,1.0825]
AA	34	0.9963	[0.9891,1.0036]	0.9923	[0.9730,1.0119]	1.0064	[0.9284,1.0909]	0.9645	[0.8306,1.1200]
GA+AA	150	0.996	[0.9890,1.0025]	0.9901	[0.9727,1.0078]	0.9816	[0.9210,1.0462]	0.9647	[0.8635,1.0779]

When the 7 day exposure associations were examined, a similar relationship with the 'A' allele and observed, but this time in the Asian children, **Table 5.10**. In the Black childen when this exposure interval was considered the 'GG' genotype was associated with increased excretion of 8-isoP, again consistent with the protective action of the 'A' allele. No associations were evident with the annual exposure estimates.

5.3.4.3 Interaction between Genotype and Biomarkers of Traffic Exposure

No significant modifying effects were noted with any of the studied SNPs in SOD3 on the relationship between 24 hour (**Appendix J, Tables J6** and **J7**) and 7-day (**Appendix J, Tables J10** and **J11**) exposure attributions and the urinary biomarkers of oxidative stress. In addition no influence of any of the studied SNPs, in the three examine genes, were noted on the relationship between the urinary markers and the annual exposure attributions (**Appendix J**, **Tables J12-J15**).

The three genes related to aspects of the cellular response to oxidative stress were not a proiri selected for consideration for a modifying effect on the association between modeled air pollutant exposures and urinary traffic biomarkers. Here the associations are shown only for completeness. The impact of genetic variants in SOD3, Nrf2 and GLCM on the association between acute (24h), sub-chronic (7 day) and chronic (annual) exposures to the modeled pollutants and the measured urinary Cu and Ni concentrations are shown in **Appendix J**, **Tables J16** to **J27**. Overall, following correction for multiple comparisons, no significant interactions were apparent between the examined genotypes and the association between pollutant exposure and urinary Cu and Ni. In the absence of any significant associations, I elected not to over interpret the associations reported at nominal significance levels.

5.4 DISCUSSION

This study represents a comprehensive investigation of urinary biomarkers of response (oxidative stress) and exposure (metals) associated with urban gaseous and particulate air pollution. I hypothesized that children with high modeled exposures to pollution would show significantly elevated levels of urinary biomarkers of oxidative stress, in addition to increased urinary excretion of metals reflecting traffic emissions from exhaust and/or vehicular abrasion. In addition, I examined whether the urinary concentration of these biomarkers would be influenced by variation in genes related to protection (GCLM and SOD3) and adaption to oxidative stress (Nrf2).

In the crude analysis I observed evidence that urinary concentrations of 8-oxo-2'-deoxyguanosine (8-oxodG) were positively associated with 24h exposures to NOx, NO_2, $PM_{2.5}$ and PM_{10}, whilst the associations with 8-isoprostane were weakly negative for the same pollutants across all of the studied exposure durations. Following adjustment for the selected co-variants and multiple testing for the exposure attributions investigated, only the increase in 8-oxodG in relation to the 24h $PM_{2.5}$ and PM_{10} exposures and the negative associations between 8-isoprostane and 7-day average NOx and NO_2 remained significant. The concentration of urinary 8-oxodG in relation to the acute to sub-chronic exposures to PM_{10} appeared to be increased in individuals carrying the 'C' allele in Nrf2, rs2364723. Urinary 8-isoprostane concentrations in relation to sub-chronic exposures to NO_2, were also found to be increased in individuals possessing the 'C' allele in Nrf2, rs7557529, in the combined model; whilst they were decrease in the population with the 'A' allele in GLCM, rs3170633 with sub-chronic exposures to $PM_{2.5}$. No associations were observed between the markers of oxidative stress and the long term exposure attributions suggesting that these markers are largely influenced by short periods of accrued exposure.

I also observed increased metal concentrations (Ba, Mn, Sn and Ni) in spot urine samples from children exposed to ETS, which was associated with increased urinary concentrations of 8-oxodG and 8-isoprostane. When the urinary metal concentrations

were examined in relation to the acute to chronic exposure attributions, it was notable that the correlations were broadly negative with the shorter term attributions, but became significantly positive for Cu, Fe, Ba. Mn, Cd, As, Ni, and particularly Cr and V when annual exposure estimates were applied, particularly for $PM_{2.5}$. More detailed analysis of the urinary metal data was restricted to Cu, (being indicative of an abrasion metal source) and Ni (reflecting long range oil combustion) (Gerlofs-Nijland et al. 2007; Thorpe and Harrison 2008; Godri et al. 2011). These analyses confirmed the negative association between urinary Cu and sub-chronic exposures to NOx and acute exposure to $PM_{2.5}$. A similar negative relationship was also noted between urinary Ni and sub-chronic exposures to PM_{10} and $PM_{2.5}$, after adjustment for all of the studied covariants.

In the present investigation urinary concentrations of 8-oxodG and 8-isoprostane were employed as markers of systemic oxidative stress as urinary concentrations have previously been shown to be increased in pulmonary conditions with an acknowledged component of oxidative stress (Igishi et al. 2003; Wu et al. 2007), in response to cigarette smoke (Campos et al. 2011; Mesaros et al. 2012; Chiang et al. 2012), occupational and environmental exposures to PAHs (Andreoli et al. 2012; Lee et al. 2012; Fan et al. 2012), and indoor/outdoor pollution (Chuang et al. 2003; Lai et al. 2005; Barbato et al. 2010; Ren et al. 2011; Huang et al. 2012; Beerappa et al. 2013; Commodore et al. 2013). 8-oxodG is a major water-soluble oxidized nucleoside formed by free radicals, which is cleaved from DNA and excreted in urine. The determination of urinary 8-oxodG concentrations therefore represents a non-invasive, technically simple procedure, which provides a measure of both oxidative damage and DNA repair capacity (Cooke et al. 2002). In vitro studies have demonstrated that constituents of diesel exhaust particulates and ambient PM have the ability to produce 8-oxodG lesions in DNA through the imposition of cellular oxidative stress (Arimoto et. al. 1999; Knaapen et. al. 2002). F2-isoprostanes, including 8-isoprostane are considered a good standard biomarker of oxidative stress in vivo as they are chemically stable and appear independent of dietary lipid intake (Il'yasova et al. 2012). Concentrations of 8-isoprostane have been shown to be elevated in exhaled breath condensate of subjects

with respiratory disease (Kharitonov and Barnes 2004) and after exposure to inhaled irritants, including ambient air pollution (Huang et al. 2012; Patel et al. 2013). In addition, urinary concentrations of this stable lipid peroxidation marker have been shown to be increased in populations exposed to traffic fumes (Rossner et al. 2007), particulate matter derived from wood combustion (Barregard et al. 2006) and exposure to cigarette smoke (Noakes et al. 2007; Seet et al. 2011).

Previous studies have shown significant differences in the level of oxidative DNA damage between ethnicities, with oxidative DNA damage reported to be lower in African-Americans compared with white subjects by Huang et al. (2000). Similarly, in the present study urinary 8-oxodG and 8-isoprostane concentrations varied significantly between the studied ethnicities, but did not accord with the previous observations, with the Asian children having the highest concentrations of these two biomarkers. This variation may in part be related to the children's diet, as diets rich in antioxidants may protect against oxidative stress (Kelly 2004). However, antioxidant supplementation studies have not definitively shown evidence of reduced urinary 8-oxodG levels (Huang et al. 2002) and it is therefore likely that the difference observed reflects either genetic difference in DNA repair pathways, or exposure to exogenous pro-oxidants. Gender has also been previously reported as an important confounder for biomarkers of oxidative stress, mainly due to differences in metabolic rates amongst males and females (Loft and Poulsen 1998). In the present study, I observed significant differences in the concentrations of 8-oxodG and 8-isoprostane between the males and female subjects in the crude analysis, but these differences were lost in the fully adjusted model.

Exposure to tobacco smoke (ETS) was also associated with increased concentrations of 8-oxodG and 8-isoprostane in the crude analysis, with the difference in the later remaining in the adjusted model. Children exposed to ETS also showed elevated metal concentrations, with increased urinary concentrations of Ba, Mn, Ni and Sn. This concurred with previous reports of elevated metal concentrations in passive and active smokers (Richter et al. 2009). This is plausible as many of the metals shown to be elevated in the children's urine are known to be present in the cigarette smoke aerosol (Chiba and Masironi 1992). Previous work has also shown that the metal

content of cigarette smoke PM is responsible for the induction of oxidative stress (Ghio et al. 2008; Pappas 2011). It was notable however, that Fe and Cu, which are known to be abundant in ETS particulate, were not elevated in the urine implying their bioaccumulation in vivo.

Urban particulate matter also contains a variety of transition metals adsorbed on their surfaces which have potential to produce reactive oxygen species in vivo resulting in oxidative stress (Costa and Dreher 1997; Li et al. 2003; Kelly et al. 2011). In this study I therefore performed exploratory analysis to examine whether there were associations between the modeled pollutant attributions and urinary metals that would support their use as biomarkers of either, non-tail pipe (Cu, Sb, Ba, Fe etc.), or tail pipe (Cr, Ni, V etc.) sources (Thorpe and Harrison 2008). In addition, as many of the studied metals are either redox catalysts (Fe, Cd, Ni, V, Cr, Mn and Cu) (Valko et al. 2005), or can deplete cellular antioxidant pools through the formation of GSH mercaptides (Cd, Cu, Zn, As and Pb) (Ballatori 1994), they also provide a mechanistic basis for the induction of systemic oxidative stress in populations with high exposures to traffic pollutants. The potential of metals and particularly Fe to influence air pollutant, health associations can be inferred from the series of papers arising out of the Normative Aging study, in which SNPs in hemochromatosis protein (HFE) has been shown to modify the relationship between $PM_{2.5}$ and traffic exposures and heart rate variability (Park et al. 2006) and cognitive decline (Power et al. 2013).

Of the metals examined detailed analysis was restricted to just Cu and Ni, as being reflective of abrasion and tail pipe / oil combustion sources respectively. Whilst there was some evidence of weakly positive correlations with both Cu and Ni with annual $PM_{2.5}$ and PM_{10} exposure attributions, the relationship over the shorter term exposure intervals was predominately negative; a pattern seen with the majority of the trace elements examined. After adjustment for the covariants in the full model only the negative associations with sub-chronic exposures remained significant, though the basis for this inverse relationship was not clear. These data therefore do not support the contention that urinary metals can be used as biomarkers of traffic exposure. Whilst many trace elements are excreted in the urine, Cu, Fe, Zn and Mn (all potentially useful

biomarkers of vehicular abrasion processes) are predominately excreted through faecal route, which may limit their usefulness as simple exposure biomarkers (Iyengar 1998).

The positive association observed in the present study between urinary 8-oxodG concentrations and 24h exposures $PM_{2.5}$ and $PM_{10,}$ broadly fits with the increase in 8-oxodG reported in aged subjects by Ren et al. (2011), though in this later study the associations with the pollutant metrics examined (significant for $PM_{2.5}$, NO_2, ozone, sulphate and organic carbon, but not black carbon, CO or particle number concentration) were based on 10-21 day moving average exposures, similar to the sub-chronic attributions used in the present analysis. The associations observed were argued to reflect evidence that the observed DNA oxidation was related to secondary, rather than primary air pollutants. A subsequent analysis of the same data demonstrated that the observed associations between sulphate and organic carbon exposures with urinary 8-oxodG concentrations were modified by polymorphisms in GSTP1, GSTM1 and catalase, though the direction of the modification varied with the polymorphisms examined and the mechanistic basis for the observed effect modifications remained oblique (Ren et al. 2010). A study performed on 894 (6-10 year old) children in the Czech Republic also reported positive associations between urinary 8-oxodG concentrations and ambient PM_{10}, $PM_{2.5}$, particulate PAH and benzo(a)pyrene exposures considering a variety of exposure durations, 3-7 days (Svecova et al. 2009). In addition urinary 8-oxodG concentrations have been found to be significantly elevated in children living in Bangkok compared with rural areas of Thailand (Buthbumrung et al. 2008). Increased urinary 8-oxodG concentrations have also been reported in Taiwanese traffic conductors compared with office workers, with a positive association with urinary 1-hydroxypyrene-glucuronide, a biomarker of traffic derived PAHs (Huang et al. 2012), as well as in long-distance bus drivers (Han et al. 2010). Other studies have failed to demonstrate a simple relationship between the exposure to high ambient pollutant concentrations and urinary 8-oxodG levels, even in the presence of increased urinary PAH metabolites (Fan et al. 2012; Rossner et al. 2013). A comparison of high versus low pollution communities in the Czech Republic showed no association between $PM_{2.5}$ exposures (both short and multi-day) and urinary 8-oxodG

concentrations, despite evidence of associations with other markers of systemic oxidative stress (plasma 15_{F2t}-isoprostane and protein carbonyls) (Rossner et al. 2013).

In the present study, the increase in urinary 8-oxodG concentrations associated with the previous 24 hour exposure to $PM_{2.5}$ and PM_{10} occurred in parallel to a decrease in excreted 8-isoprostane, in relation to 7 days NOx and NO_2 exposures. This disparity in part suggests that different components of the urban aerosol are acting through separate pathways in response to different pollutant triggers. As the presence of 8-oxodG in the urine not only reflects damage to DNA and oxidation of the nucleotide pool (Sorensen et al. 2003), but also the excision repair capacity of the organism (Cooke et al. 2003), its increased concentrations implies the activation of the base excision repair pathway, i.e. both oxidative damage and cellular adaptation. Given that there is evidence that Nrf2 plays a role in 8-oxoguanine DNA glycosylase (OGG1) expression (Singh et al. 2013), it is therefore possible that the decrease in urinary 8-isoprostane might reflect the up-regulation of antioxidant genes under the regulation of this transcription factor. Hence, whilst the associations are in the opposite direction, they might reflect the same ongoing process.

As the formation of DNA base and lipid oxidation markers reflects the presence of oxidative stress, I investigated whether genetic variants in the antioxidant enzyme extra-cellular superoxide dismutase (SOD3), as well as regulators of the antioxidant response (GCLM and Nrf2) would potentially modify any underlying association between pollutant exposure and systemic oxidative stress. Nrf2, a basic leucine zipper transcription factor plays an essential role in the regulation of cyto-protective responses, through the induction of a series of antioxidant and detoxification enzymes that neutralize ROS and other electrophiles, including glutamate cysteine ligase regulatory subunit (-GCS) and superoxide dismutase (SOD) (Cho and Kleeberger 2010). It has been hypothesized that functional polymorphism in Nrf2 may be associated either with a change in the expression of basal Nrf2 or may cause an inability of Nrf2 to translocate from cytoplasm to its binding sites in nucleus. This would then result in attenuation of ARE (antioxidant response element)-related gene transcription exposing the individuals to a higher risk of oxidative stress. SNPs in the Nrf2 promoter region have previously

been shown to influence transcriptional activity and Nrf2 expression in vitro (Marzec et al. 2007). Several recent studies have investigated genetic polymorphisms in Nrf2 in relation to respiratory disease susceptibility, either singularly or associated with other polymorphisms, related to acute lung injury (Marzec et al. 2007), lung function decline due to smoking exposure (Siedlinski et al. 2009; Henderson et al. 2010; Masuko et al. 2011) and asthma in children exposed to highly polluted environments (Ungvári et al. 2012).

In the present study, urinary 8-oxodG concentrations in relation to sub-chronic exposures to PM_{10} and $PM_{2.5}$ appeared to be increased in individuals carrying the 'C' allele in Nrf2, rs2364723 (G/C). Previous studies have shown that heterozygous (GC) of rs2364723 is negatively associated with the FEV_1 in a large Caucasian cohort (Siedlinski et al. 2009). Similarly, in a small study examining Asian lung cancer smokers (Sasaki et al. 2013), mean FEV_1 was found to be significantly lower in the groups homozygous for CC compared to wild type GG.

Urinary 8-isoprostane concentrations in relation to sub-chronic exposures to NOx and NO_2, were also found to be increased in individuals possessing the 'C' allele in Nrf2, rs7557529 (T/C), in the combined model. This is a novel interaction, with no previous reports of effect modification with this SNP. In contrast, possession of the 'A' allele in the combined model for GLCM, rs3170633 (G/A) conferred protection against urinary 8-isoprostane in relation to sub-chronic $PM_{2.5}$ exposures. Glutamate-cysteine ligase (GCL), the first rate-limiting enzyme, catalyzes synthesis of GSH and is induced by Nrf2 (Jyrkkänen et al. 2008). Polymorphisms in GCLM has been associated with lower promoter activity, significantly lower plasma GSH concentration (Nakamura et al. 2002) and increased susceptibility to oxidative stress (Yang et al. 2002). GCLM has been investigated by various researchers as a protective candidate gene in lung function (Siedlinski et al. 2009), COPD (Chappell et al. 2008) and susceptibility to schizophrenia (Tosic et al. 2006).

Several studies have shown a role for polymorphisms in SOD3 in mediating the susceptibility of the lung to oxidative stress and lung function changes in respiratory disease (Dahl et al. 2008; Ganguly et al. 2009), COPD (Dahl et al. 2008; Juul et al. 2006; Young et al. 2006), acute lung injury (Arcaroli et al. 2009), cancer risk (Kang et al. 2007) and risk factor for cardiovascular disease (Naganuma et al. 2008a; 2008b). Extracellular SOD is expressed abundantly in the lungs, within the respiratory tract lining fluids, at the respiratory epithelium and vascular endothelium (Oury et al. 1994, 1996). The SOD3 gene has been localized to chromosome 4 and consists of three exons and two introns. In the present study, we studied five polymorphisms of SOD3. One was excluded due to deviation from Hardy-Weinberg (rs2284659). Two other SNPs located in intron region (rs8192287 G/T, rs8192288 G/T) and another two found to be in promoter region (rs699473 C/T, rs13306703 C/T). I observed no significant effect modification with any of these SNPs following adjustment for multiple comparisons.

In summary, these data provide evidence that short term PM exposures in children are associated with increased DNA oxidation. As an inverse relationship was observed with the lipid oxidation biomarker 8-isoprostane, this strongly argues that this is not simplistically related to pollutant induced oxidative stress, but rather the induction of a robust DNA repair response. One might therefore conjecture that the absence of increased urinary 8-oxodG following short term exposures, is the more detrimental observation, implying that oxidatively damaged nucleotides are not being cleaved from DNA, increasing potential coding errors and the risk of mutations. Thus SNPs in genes promoting increased 8-oxodG concentrations, such as rs7557529 might actually be considered beneficial. To fully substantiate this view it would be necessary to examine a broader range of oxidative biomarkers, both in plasma and urine, and to examine associations over shorter durations, which might capture associations prior to the induction of Nrf2 mediated adaptations. In this chapter, I also examined a panel of metals in urine as potential biomarkers of exposure to traffic pollution. At this time, the results of this analysis do not appear promising, though ETS metal signatures were detected in the children's urine.

Chapter 6

Summary and Conclusion

In this thesis, I examined the impact of urban, specifically traffic-related pollution, on the respiratory health of children living within East London. The study was conducted over the first three years of the lifetime of London's Low Emission Zone, and as such represents a baseline study to evaluate the later impacts of the scheme, notwithstanding the fact that some year-on-year improvements in air quality might have been expected due to modernization of the vehicle fleet over this period.

The key findings arising from this study were as follows:

- Over the first three years of the LEZ lifetime I observed no evidence of measureable improvements in urban air quality.
- Significant associations were observed between annual pollutant exposures (oxides of nitrogen, NO_2, $PM_{2.5}$ and PM_{10}) with current rhinitis and FVC in children living within Tower Hamlets and Hackney.
- Through the development of a temporal scaling tool (NOWCAST) and its application to our annual dispersion models I was able dissect out the short, versus long-term influence of air pollution on the health endpoints examined. As the reduction in FVC was only associated with annual and not short-term (1-day to 1-week) exposures, I have argued that this provides evidence of reduced lung growth.
- I observed evidence that polymorphisms in CYP1A1 modify acute and sub-chronic effect of air pollution on lung function. One protective (rs2198843) and one sensitizing SNP (rs2606345) were identified. Notably, these associations were most marked within the black and Asian children studied.

- Acute 24h exposures to PM_{10} and $PM_{2.5}$ were associated with increased, excretion of urinary 8-oxodG, with the effect most marked in white children heterozygous for the 'C' allele in rs2364723 (Nrf2). In contrast, urinary concentrations of 8-isoP were decreased in relation to 24h NO_2 and NOx exposures.
- Urinary metals, Ni and Cu, reflecting tail pipe oil combustion and vehicular abrasion sources, were investigated as biomarkers of traffic exposure. This exploratory analysis demonstrated some evidence of weakly positive correlations with both Cu and Ni and annual $PM_{2.5}$ and PM_{10} exposure attributions. However over shorter term exposure intervals the relationship was predominately negative. At this time therefore these data do not support the of these urinary metals as traffic biomarkers.

Study Strengths and Limitations:

Whilst initially conceived to study the period running up to and after the introduction of the LEZ, due to funding issues and the time required to obtain ethics this study was only able start after the first two phases of the LEZ had already been introduced. It would clearly have been better to have had a run in period, as it now appears that by the beginning of 2008 most of the HGV fleet targeted by the tighter emission standards had already upgraded (personal communication, Dr Ben Barratt). Nethertheless it was hoped (and predicted by pre and post scheme modelling exercises) that phase 3 of the LEZ would produce a sufficient improvement in air quality to detect a signal in respiratory endpoints. The impact of the political decision to delay this phase had significant ramifications on this study and its power to detect an air pollution signal. The LEZ however was not the only instrument acting over this period that had the potential to improve urban air quality, as natural fleet turnover and the increased proportion of newer vehicles with EURO IV and V engines should also have produced air quality improvements. Hence throughout all of the analysis we have included year in our models, but have failed to see any annual improvements in respiratory symptoms or

lung function. In terms of lung function, especially in relation to restrictive patterns in children that might reflect deficits in optimal growth, a critical question is how long would be required for measureable decrements to manifest. In the original rationale for the study it was thought that this would reflect the proportion of life each cross sectional panel of children had lived within the zone. For year 1, between 8-13 months depending on the date the schools were visited; for year 2, 20-25 months and so on. Thus for a cohort aged between 8-9 years, the proportion of their life spent within the zone would be approximately , 7.4-13.5% in year 1, raising to 27.8-38.5% in year 3. By the final year of the study, assuming children have remained within the study area for the full 6 year duration, the proportion of their lifetime lived within this traffic management area will be between 63.9-76.0%. Given the rate of lung growth at the age selected and the accrued effects of living within the LEZ, it was reasonable to expect signals related to long term effects on lung function.

The current study was focused in an area of East London with very poor air quality and examined respiratory health in a deprived and multi-ethnic cohort of children. In particular, this is the first time a central Asian population of children have been examined in this regard. It employed a validated high resolution dispersion model to provide estimates of annual exposures, as well as scaling factors to provide shorter term exposure attributions based on measurements made within and surrounding the study area, on the same spatial scale. It was thus able to discriminate between short and long term impacts of air pollution on lung function and systemic oxidative stress. This permitted inferences to be made on lung growth, which otherwise would not have been possible in a cross sectional study. Clearly a longitudinal follow up addressing lung function in these children would be required to confirm this inference. In term of lung function, this is the only study to date that has examined lung function post bronchodilator in relation to pollutant exposures and made use of the same trained personal throughout the three years of the study, both for the actual measurements and data quality control. In addition, each child also provided a questionnaire on respiratory symptoms in the past year, as well as FeNO as a measure of allergic airway inflammation (Rolla et al. 2007). Urine samples were also provided for the assessment

of biomarkers of pollutant exposure and systemic oxidative stress, as well as saliva and mouth swabs for the determination of vitamin D status and collection of DNA for the genotyping component of the project.

Whilst examining the ethnically diverse population within central London was one of the strengths of the current study, it brings with it inherent difficulties in the analysis and interpretation of the genetic data. Population stratification is capable of producing spurious findings in gene association studies, especially when the study population consists of two or more ancestral populations. In addition, differences in the admixture proportion of individuals may also produce false positives if the outcome is sensitive to the ancestry proportion. Although, I adjusted for self-reported ethnicity in the regression models employed to control for the effect of confounding by population stratification, as the allele frequencies for some of the SNPs differ across ethnic groups, the use of ancestry markers to control population structure might have improved the analysis. However, my attempt to correct the population structure using ancestry markers was not successful at this point, though based only on a limited number of random SNPs. In addition, using these ancestry markers did not substantially modify the results, though they were only really effective at separating out the black from Eurasian populations. Better separation might have been achieved using a more comprehensive set of ancestral informative markers as has been previously been shown (Winkler et al. 2010). In this thesis I have therefore also analysed the associations identified as significant within the whole population separately for the three main ethnic groups, with reference to that population's major allele for the SNPs investigated. Overall these analyses produced results coherent with that observed for the whole population, though they did demonstrate that some polymorphisms had their effect predominately in only one of the ethnicities.

In the present investigations defined panels of polymorphisms in genes were selected for each set of endpoints, based either on pre-published evidence of associations with air pollutant exposures (such as with the glutathione S transferases), or a prior hypothesis based on the known, or purported mechanisms of air pollutant toxicity (CYP1A1 and the AhR). This approach was taken to avoid the high statistical

penalty inherent in examining large numbers of SNPs. The human genome has been shown to contain a haplotype block structure, consisting of various distinct blocks within which SNPs are strongly related. Within each block, a number of SNPs called 'tag SNPs' can be utilized to differentiate a large fraction of the haplotypes. These tagging SNPs are very useful in the association studies as they eliminates the need of genotyping all SNPs within the same block and hence eliminate the high statistical penalty of examining multiple SNPs. Haplotype-based association studies are therefore preferred over the genotype-by-genotype studies in the genes, containing multiple susceptible alleles as they are thought to be more effective (Zhang 2002). Gauderman et al. (2007) reported that using multiple SNPs that capture underlying genetic architecture for the disease/outcome are more powerful at detecting gene-environment interactions. Going forward such a haplotype block approach should be favoured, as the size of the current cohort, even after inclusion of the children recruited across years 4-6 of the study (approx. 2,200) remains too small for a successful GWAS. Of the panels of SNPs examined, had funds permitted, it would have been good to have examined polymorphisms in genes related to the base excision repair pathway (Wilson et al. 2011) to more fully define the nature of increased urinary concentrations of 8-oxodG observed. As was inferred in the discussion to chapter 5 it remains unclear whether this increase should be viewed as an adverse response, or evidence of the appropriate induction of DNA repair at this time. Finally, whilst there is a considerable evidence base relating genetic variation in phase I and II xenobiotic and antioxidant genes to the relationship between respiratory and cardiovascular responses to air pollutant exposures (Minelli et al. 2011), there are considerably fewer studies addressing key inflammatory genes. This literature was recently reviewed (Vawda et al. 2013) and reported evidence supporting a role for both toll-like receptor 4 and transforming growth factor β1 in mediating the association between lung function and air pollutant exposures. Further focus on SNPs within these genes therefore warrants attention.

Future and Ongoing Work

The use of models to assign pollution exposures is difficult, mainly due to the fact that the models assumed that subjects reside at the residential address 24h a day for 365 days a year, which is implausible. Whilst we attempted to evaluate home, versus home plus school exposures, we made no attempt to model exposures within the home, or to take into account the children's journey to and from school. As such, similar to many other studies, there is a considerable risk of exposure misclassification. Indoor exposures to pollutants are also likely to be high in households using gas cookers, which was not controlled for in the current study until years 4-6. Kornartit et al. (2010) found that personal exposure to NO_2 was strongly correlated to indoor NO_2 concentrations. Future work is required to establish the magnitude of this exposure misclassification comparing personal to modelled exposures in large cohorts, though it is likely this is smaller in children than adults, due to the more restricted geographical area they move through. It should be stated that the number of children within our cohort who did not walk to school was very low, <10%, suggesting that in vehicle exposures were unlikely to confound our results. Much of this uncertainty could be dealt with if good biomarkers of traffic exposure could be identified. In this study, I examined urinary metals, but from these preliminary analyses, these do not appear to be useful in characterizing primary vehicle exposures. As a further extension to this study, metabolomics analyses are currently ongoing in an attempt to isolate an exposure biomarker using both NMR and LC-MS based approaches, using urine samples from the 400 children with the highest and lowest NOx exposures within our cohort. Should such a biomarker be identified it would greatly improve our understanding of the contribution of traffic, specifically diesel exhaust, to the health endpoints examined.

Of the routinely measured parameters in this report, I have only reported on respiratory symptoms, lung function and the urinary biomarkers. Though FeNO measurements were made in all subjects, these were often below the detection limit of the NIOX instrument used and at the time of analysis these data would not converge in the models employed, either as a continuous or categorical variable. Despite this limitation the FeNO measurements were found to align well with the reported

symptoms of asthma, hay fever and rhinitis (data not shown) suggesting that they were discriminating between individuals with and without allergic inflammation. As with the other major endpoints employed in this study a panel of SNPs were selected a priori to interrogate any observed associations between FeNO concentration and air pollutant attributions, focused on polymorphisms in iNOS, arginase 1 and 2 based on earlier work by Salam et al. (2011, 2012) and as such analysis of this data remains a priority moving forward.

The determination of vitamin D in saliva was not possible after a considerable period spent developing the pre-existing analysis methodology (Martineau et al. 2011), despite previous reports (Higashi et al. 2008). The saliva was therefore used for the determination of markers of innate IgA, (Ewing et al. 2010) and allergic inflammation, IgE (Mimura et al. 2010), as well as the assessment of cortisol, as a measure of activity of the hypothalamic-pituitary-adrenal axis (Stenius et al. 2011). In a small subset of children, induced sputum was also performed on a second visit to examine particle uptake according to the method of Kulkarni et al. (2006) as an alternative biomarker of exposure.

The LEZ school study is ongoing due to the delay in the implementation of phases III and IV, and is due for completion in March 2014. At this time, the full effectiveness of the scheme will be evaluated, with an increased power than was available in this preliminary analysis.

REFERENCES

1. Abnet CC, Fagundes RB, Strickland PT, Kamangar F, Roth MJ, Taylor PR, Dawsey SM. 2007. The influence of genetic polymorphisms in Ahr, CYP1A1, CYP1A2, CYP1B1, GST M1, GST T1 and UGT1A1 on urine 1-hydroxypyrene glucuronide concentrations in healthy subjects from Rio Grande do Sul, Brazil. Carcinogenesis. 28(1):112-7.
2. Adonis M, Martínez V, Riquelme R, Ancic P, González G, Tapia R, Castro M, Lucas D, Berthou F, Gil L. 2003. Susceptibility and exposure biomarkers in people exposed to PAHs from diesel exhaust. Toxicology Letters 144(1):3–15.
3. Agency for Toxic Substances and Disease Registry (ATSDR). 1997. Toxicological Profile for Titanium Tetrachloride. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.
4. Agency for Toxic Substances and Disease Registry (ATSDR). 2005. Toxicological profile for Tungsten. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.
5. Agency for Toxic Substances and Disease Registry (ATSDR). 2012. Toxicological Profile for Manganese. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.
6. Aguilera I, Daponte A, Gil F, Hernández AF, Godoy P, Pla A, Ramos JL; DASAHU group, Daponte A, Aguilera I, Fernández-Ajuria A, Toro S, Martín-Olmedo P, Lacasaña M, Mayoral JM, Pla A, Gil F, Hernández A, Villanueva E, Rodrigo L, de Santiago E, López O, Ramos JL, Godoy P, Sánchez-Parra F. 2010. Urinary levels of arsenic and heavy metals in children and adolescents living in the industrialised area of Ria of Huelva (SW Spain). Environ Int. 36(6):563-9.
7. Alexeeff SE, Litonjua AA, Wright RO, Baccarelli A, Suh H, Sparrow D, Vokonas PS, Schwartz J. 2008. Ozone exposure, antioxidant genes, and lung function in an elderly cohort: VA normative aging study. Occupational and Environmental Medicine 65(11):736–742.
8. Alexis NE, Zhou H, Lay JC, Harris B, Hernandez ML, Lu TS, Bromberg PA, Diaz-Sanchez D, Devlin RB, Kleeberger SR, Peden DB. 2009. The glutathione-S-transferase Mu 1 null genotype modulates ozone-induced airway inflammation in human subjects. J Allergy Clin Immunol. 124:1222–1228.

9. Anderson HR, Ruggles R, Pandey KD, Kapetanakis V, Brunekreef B, Lai CK, et al. 2010. Ambient particulate pollution and the world-wide prevalence of asthma, rhinoconjunctivitis and eczema in children: Phase One of the International Study of Asthma and Allergies in Childhood (ISAAC). Occup Environ Med 67:293-300.
10. Anderson RA. Chromium, glucose tolerance, and diabetes. 192. Biol Trace Elem Res. 32:19-24.
11. Andreoli R, Protano C, Manini P, De Palma G, Goldoni M, Petyx M, Rondinone BM, Vitali M, Mutti A. 2012. Association between environmental exposure to benzene and oxidative damage to nucleic acids in children. Med Lav. 103(5):324-37.
12. Andrysík Z, Vondráček J, Marvanová S, Ciganek M, Neča J, Pěnčíková K, Mahadevan B, Topinka J, Baird WM, Kozubík A, Machala M. 2011. Activation of the aryl hydrocarbon receptor is the major toxic mode of action of an organic extract of a reference urban dust particulate matter mixture: the role of polycyclic aromatic hydrocarbons. Mutation Research 714(1-2):53–62.
13. Anttila S, Lei XD, Elovaara E, Karjalainen A, Sun W, Vainio H, Hankinson O. 2000. An uncommon phenotype of poor inducibility of CYP1A1 in human lung is not ascribable to polymorphisms in the AHR, ARNT, or CYP1A1 genes. Pharmacogenetics 10: 741–751.
14. Apostoli P, Neri G, Lucas D, Manno M, Berthou F. 2003. Influence of genetic polymorphisms of CYP1A1 and GSTM1 on the urinary levels of 1-hydroxypyrene. Toxicology Letters 144(1):27–34.
15. Arcaroli JJ, Hokanson JE, Abraham E, Geraci M, Murphy JR, Bowler RP, Dinarello CA, Silveira L, Sankoff J, Heyland D, Wischmeyer P, Crapo JD. 2009. Extracellular Superoxide Dismutase Haplotypes Are Associated with Acute Lung Injury and Mortality. American Journal of Respiratory and Critical Care Medicine 179: 105–112.
16. Arcus-Arth A, Blaisdell RJ. 2007. Statistical distributions of daily breathing rates for narrow age groups of infants and children. Risk Analysis 27(1):97–110.
17. Arimoto T, Yoshikawa T, Takano H, Kohno M. 1999. Generation of reactive oxygen species and 8-hydroxy-2'-deoxyguanosine formation from diesel exhaust particle components in L1210 cells. Japanese Journal of Pharmacology 80: 49–54.
18. Arisawa T, Tahara T, Shibata T, Nagasaka M, Nakamura M, Kamiya Y, Fujita H, Yoshioka D, Okubo M, Hirata I, Nakano H. 2008. Nrf2 gene promoter polymorphism and gastric carcinogenesis. Hepatogastroenterology 55: 750–754.
19. Arneson N, Hughes S, Houlston R, Done S. 2008. GenomePlex Whole-Genome Amplification. CSH Protocols 2008 pdb.prot4920.

20. Arnich N, Sirot V, Rivière G, Jean J, Noël L, Guérin T, Leblanc JC. 2012. Dietary exposure to trace elements and health risk assessment in the 2nd French Total Diet Study. Food Chem Toxicol. 50(7):2432-49.
21. Ashley F, Kannel WB, Sorlie PD, Masson R. 1975. Pulmonary function: Relation to aging, cigarette habit, and mortality; The Framingham Study. Ann Intern Med 82:739–745.
22. Audi, G, Wapstra GAH. The percent natural abundance data is from the 1997 report of the IUPAC Subcommittee for Isotopic. Nucl. Phys A. 1993, 565, 1-65 and G. Audi, A. H. Wapstra Nucl. Phys A. 1995, 595, 409-480
23. Austin JB, Kaur B, Anderson HR, Burr M, Harkins LS, Strachan DP, et al. 1999. Hay fever, eczema, and wheeze: a nationwide UK study (ISAAC, international study of asthma and allergies in childhood). Arch Dis Child 81:225-230.
24. Avogbe PH, Ayi-Fanou L, Autrup H, Loft S, Fayomi B, Sanni A, Vinzents P, Møller P. 2005. Ultrafine particulate matter and high-level benzene urban air pollution in relation to oxidative DNA damage. Carcinogenesis 26(3):613–20.
25. Avol EL, Gauderman WJ, Tan SM, London SJ, Peters JM. 2001. Respiratory effects of relocating to areas of differing air pollution levels. Am J Respir Crit Care Med 164:2067–2072.
26. Ayres JG, Borm P, Cassee FR, Castranova V, Donaldson K, Ghio A, et al. 2008. Evaluating the toxicity of airborne particulate matter and nanoparticles by measuring oxidative stress potential--a workshop report and consensus statement. Inhalation toxicology 20(1):75–99.
27. Badmaev V, Prakash S, Majeed M. 1999. Vanadium: a review of its potential role in the fight against diabetes. Journal of Alternative and Complementary Medicine 5: 273–291.
28. Bagryantseva Y, Novotna B, Rossner P, Chvatalova I, Milcova A, Svecova V, Lnenickova Z, Solansky I, Sram RJ. 2010. Oxidative damage to biological macromolecules in Prague bus drivers and garagemen: impact of air pollution and genetic polymorphisms. Toxicology Letters 199: 60–68.
29. Balaji L, Singh KB, Bhaskar LVKS. 2012. CYP1A1 genotypes and haplotypes and risk of oral cancer: A case-control study in South Indians. Genetics and Molecular Biology 35(2):407–12.
30. Ballatori N. 1994. Glutathione mercaptides as transport forms of metals. Adv Pharmacol. 27:271-98.

31. Barbato DL, Tomei G, Tomei F, Sancini A. 2010. Traffic air pollution and oxidatively generated DNA damage: can urinary 8-oxo-7,8-dihydro-2-deoxiguanosine be considered a good biomarker? A meta-analysis. Biomarkers 15: 538–545.
32. Barceloux DG. 1999. Nickel. J Toxicol Clin Toxicol. 37(2):239-58.
33. Barker DL, Hansen MST, Faruqi AF, Giannola D, Irsula OR, Lasken RS, Latterich M, Makarov V, Oliphant A, Pinter JH, Shen R, Sleptsova I, Ziehler W, Lai E. 2004. Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. Genome Research 14: 901–907.
34. Barraza-Villarreal A, Sunyer J, Hernandez-Cadena L, Escamilla-Nuñez MC, Sienra-Monge JJ, Ramírez-Aguilar M, Cortez-Lugo M, Holguin F, Diaz-Sánchez D, Olin AC, Romieu I. 2008. Air Pollution, Airway Inflammation, and Lung Function in a Cohort Study of Mexico City Schoolchildren. Environmental Health Perspectives 116: 832–838.
35. Barregard L, Sällsten G, Gustafson P, Andersson L, Johansson L, Basu S, Stigendal L. 2006. Experimental exposure to wood-smoke particles in healthy humans: effects on markers of inflammation, coagulation, and lipid peroxidation. Inhal Toxicol. 18(11):845-53.
36. Baruthio F. 1992. Toxic effects of chromium and its compounds. Biological Trace Element Research 32: 145–153.
37. Bateson TF, Schwartz J. 2008. Children's response to air pollutants. Journal Of Toxicology And Environmental Health Part A 71(3):238–243.
38. Beckett SM, Laughton SJ, Pozza LD, McCowage GB, Marshall G, Cohn RJ, Milne E, Ashton LJ. 2008. Buccal swabs and treated cards: methodological considerations for molecular epidemiologic studies examining pediatric populations. American Journal of Epidemiology 167(10):1260–1267.
39. Beerappa R, Venugopal D, Sen S, Ambikapathy M, Rao RH. 2013. Assessment of 8-oxo-7,8-dihydro-2'-deoxyguanosine as a marker of oxidative DNA damage in gasoline filling station attendants. Int J Occup Med Environ Health. 26(5):780-9.
40. Beevers SD, Kitwiroon N, Williams ML, Kelly FJ, Ross Anderson H, Carslaw DC. Air pollution dispersion models for human exposure predictions in London. J Expo Sci Environ Epidemiol. 2013 Feb 27. doi: 10.1038/jes.2013.6. [Epub ahead of print]
41. Behndig AF, Blomberg A, Helleday R, Duggan ST, Kelly FJ, Mudway IS. 2009. Antioxidant responses to acute ozone challenge in the healthy human airway. Inhalation Toxicology 21(11):933–42.

42. Behndig AF, Mudway IS, Brown JL, Stenfors N, Helleday R, Duggan ST, Wilson SJ, Boman C, Cassee FR, Frew AJ, Kelly FJ, Blomberg A, Sandstro T. 2006. Airway antioxidant and inflammatory responses to diesel exhaust exposure in healthy humans. European Respiratory Journal 27(2):359–365.
43. Behndig AF, Mudway IS, Brown JL, Stenfors N, Helleday R, Duggan ST, Wilson SJ, Boman C, Cassee FR, Frew AJ, Kelly FJ, Blomberg A, Sandstro T. 2006. Airway antioxidant and inflammatory responses to diesel exhaust exposure in healthy humans. European Respiratory Journal 27(2):359–365.
44. Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics 29(4):1165–1188.
45. Bennett WD, Zeman KL, Jarabek AM. 2003. Nasal contribution to breathing with exercise: effect of race and gender. Journal of Applied Physiology 95(2):497–503.
46. Bennett WD, Zeman KL, Jarabek AM. 2008. Nasal contribution to breathing and fine particle deposition in children versus adults. Journal Of Toxicology And Environmental Health Part A 71(3):227–37.
47. Bennett WD, Zeman KL. 2004. Effect of body size on breathing pattern and fine-particle deposition in children. Journal of Applied Physiology 97(3):821–6.
48. Benowitz NL. 1999. Biomarkers of environmental tobacco smoke exposure. Environmental Health Perspectives 107: 349–355.
49. Bergamaschi E, De Palma G, Mozzoni P, Vanni S, Vettori MV, Broeckaert F, Bernard A, Mutti A. 2001. Polymorphism of quinone-metabolizing enzymes and susceptibility to ozone-induced acute effects. Am J Respir Crit Care Med. 163:1426–1431.
50. Berwick M, Matullo G, Song YS, Guarrera S, Dominguez G, Orlow I, Walker M, Vineis P. 2004. Association between aryl hydrocarbon receptor genotype and survival in soft tissue sarcoma. Journal of Clinical Oncology 22(19):3997–4001.
51. Bin P, Leng S, Cheng J, Dai Y, Huang C, Pan Z, Niu Y, Duan H, Li H, Liu Q, Chen W, Zheng Y. 2008. Association of aryl hydrocarbon receptor gene polymorphisms and urinary 1-hydroxypyrene in polycyclic aromatic hydrocarbon-exposed workers. Cancer Epidemiology, Biomarkers & Prevention 17(7):1702–1708.
52. Binkova B, Chvatalova I, Lnenickova Z, Milcova A, Tulupova E, Farmer PB, Sram RJ. 2007. PAH-DNA adducts in environmentally exposed population in relation to metabolic and DNA repair gene polymorphisms. Mutation Research 620(1-2):49–61.
53. Birnboim HC. 2004. DNA yield with an Oragene self-collection kit. DNA Genotek PD–WP–001.

54. Bjermo H, Sand S, Nälsén C, Lundh T, Enghardt Barbieri H, Pearson M, Lindroos AK, Jönsson BA, Barregård L, Darnerud PO. 2013. Lead, mercury, and cadmium in blood and their relation to diet among Swedish adults. Food Chem Toxicol. 57:161-9.
55. Bobak M, Leon DA. 1992. Air pollution and infant mortality in the Czech Republic, 1986-88. Lancet 340: 1010–1014.
56. Bobak M, Leon DA. 1999. Pregnancy outcomes and outdoor air pollution: an ecological study in districts of the Czech Republic 1986-8. Occupational and Environmental Medicine 56: 539–543.
57. Bobak M, Richards M, Wadsworth M. 2001. Air pollution and birth weight in Britain in 1946. Epidemiology 12: 358–359.
58. Boettler U, Volz N, Teller N, Haupt LM, Bakuradze T, Eisenbrand G, Bytof G, Lantz I, Griffiths LR, Marko D. 2012. Induction of antioxidative Nrf2 gene transcription by coffee in humans: depending on genotype? Molecular Biology Reports 39(6):7155–7162.
59. Boezen HM, van der Zee SC, Postma DS, Vonk JM, Gerritsen J, Hoek G, et al. 1999. Effects of ambient air pollution on upper and lower respiratory symptoms and peak expiratory flow in children. Lancet 353:874-878.
60. Bonvallot V, Baeza-Squiban A, Baulig A, Brulant S, Boland S, Muzeau F, Barouki R, Marano F. 2001. Organic Compounds from Diesel Exhaust Particles Elicit a Proinflammatory Response in Human Airway Epithelial Cells and Induce Cytochrome p450 1A1 Expression. Am J Respir Cell Mol Biol. 25:515-521
61. Bonvallot V, Baulig A, Boland S, Marano F, Baeza A. 2002. Diesel exhaust particles induce an inflammatory response in airway epithelial cells: involvement of reactive oxygen species. BioFactors 16: 15–17.
62. Bowler RP, Nicks M, Tran K, Tanner G, Chang L-Y, Young SK, Worthen GS. 2004. Extracellular superoxide dismutase attenuates lipopolysaccharide-induced neutrophilic inflammation. American Journal of Respiratory Cell and Molecular Biology 31(4):432–439.
63. Bowler RP, Nicks M, Warnick K, Crapo JD. 2002. Role of extracellular superoxide dismutase in bleomycin-induced pulmonary fibrosis. American Journal of Physiology Lung Cellular and Molecular Physiology 282(4):L719–26.
64. Brauer M, Hoek G, Smit H a, Jongste JC de, Gerritsen J, Postma DS, Kerkhof M, Brunekreef B. 2007. Air pollution and development of asthma, allergy and infections in a birth cohort. The European Respiratory Journal 29(5):879–88.

65. Brauer M, Hoek G, Vliet P Van, Meliefste K, Fischer PH, Wijga A, Koopman LP, Neijens HJ, Gerritsen J, Kerkhof M, Heinrich J, Bellander T, Brunekreef B. 2002. Air Pollution from Traffic and the Development of Respiratory Infections and Asthmatic and Allergic Symptoms in Children. American Journal of Respiratory and Critical Care Medicine 166(8):1092–1098.
66. Breitner S, Stölzel M, Cyrys J, Pitz M, Wölke G, Kreyling W, Küchenhoff H, Heinrich J, Wichmann HE, Peters A. 2009. Short-Term Mortality Rates during a Decade of Improved Air Quality in Erfurt, Germany. Environmental Health Perspectives 117: 448–454.
67. Brunekreef B, Janssen N a, Hartog J de, Harssema H, Knape M, Vliet P van. 1997. Air pollution from truck traffic and lung function in children living near motorways. Epidemiology 8: 298–303.
68. Brunekreef B, Stewart AW, Anderson HR, Lai CK, Strachan DP, Pearce N. 2009. Self-reported truck traffic on the street of residence and symptoms of asthma and allergic disease: a global relationship in ISAAC phase 3. Environ Health Perspect 117:1791-1798.
69. Burr ML, Karani G, Davies B, Holmes BA, Williams KL. 2004. Effects on respiratory health of a reduction in air pollution from vehicle exhaust emissions. Occup Environ Med 61:212–218.
70. Burri PH. 1984. Fetal and postnatal development of the lung. Annual Review of Physiology 46: 617–628.
71. Burri PH. 2006. Structural aspects of postnatal lung development - alveolar formation and growth. Biology of the Neonate 89(4):313–322.
72. Buthbumrung N, Mahidol C, Navasumrit P, Promvijit J, Hunsonti P, Autrup H, Ruchirawat M. 2008. Oxidative DNA damage and influence of genetic polymorphisms among urban and rural schoolchildren exposed to benzene. Chem Biol Interact. 172(3):185-94.
73. Butler H, Ragoussis J. 2008. BeadArray-based genotyping. Methods In Molecular Biology Clifton Nj 43953–74.
74. Callan AC, Winters M, Barton C, Boyce M, Hinwood AL. 2012. Children's exposure to metals: a community-initiated study. Arch Environ Contam Toxicol. 62(4):714-22.
75. Cames M. Helmers E. 2013. Critical evaluation of the European diesel car boom - global comparison, environmental effects and various national strategies. Environmental Sciences Europe. 25:15.
76. Campos C, Guzmán R, López-Fernández E, Casado Á. 2011. Urinary biomarkers of oxidative/nitrosative stress in healthy smokers. Inhal Toxicol. 23(3):148-56.

77. Canova C, Richiardi L, Merletti F, Pentenero M, Gervasio C, Tanturri G, Garzino-Demo P, Pecorari G, Talamini R, Barzan L, Sulfaro S, Franchini G, Muzzolini C, Bordin S, Pugliese GN, Macrì E, Simonato L. 2010. Alcohol, tobacco and genetic susceptibility in relation to cancers of the upper aerodigestive tract in northern Italy. Tumori 96: 1–10.

78. Cantlay AM, Smith CA, Wallace WA, Yap PL, Lamb D, Harrison DJ. 1994. Heterogeneous expression and polymorphic genotype of glutathione S-transferases in human lung. Thorax. 49:1010–1014.

79. Carlsson LM, Jonsson J, Edlund T, Marklund SL. 1995. Mice lacking extracellular superoxide dismutase are more sensitive to hyperoxia. Proceedings of the National Academy of Sciences of the United States of America 92: 6264–6268.

80. Carlsten C, Sagoo GS, Frodsham AJ, Burke W, Higgins JP. 2008. Glutathione S-transferase M1 (GSTM1) polymorphisms and lung cancer: a literature-based systematic HuGE review and meta-analysis. Am J Epidemiol. 167:759–774.

81. Carroll WD, Lenney W, Child F, et al. 2005. Maternal glutathione S-transferase GSTP1 genotype is a specific predictor of phenotype in children with asthma. Pediatr Allergy Immunol.16 (1):32– 39

82. Carslaw DC, Beevers SD, Tate JE, Westmoreland EJ, Williams ML. 2011. Recent evidence concerning higher NOx emissions from passenger cars and light duty vehicles. Atmos Environ. 45(39):7053-63.

83. Carslaw DC. 2011. Defra Urban Model Evaluation Analysis – Phase 1. Available: http://uk-air.defra.gov.uk/library/reports?report_id=654.

84. Cascorbi I, Brockmöller J, Roots I. 1996. A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. Cancer Research 56: 4965–4969.

85. Castaño A, Sánchez-Rodríguez JE, Cañas A, Esteban M, Navarro C, Rodríguez-García AC, Arribas M, Díaz G, Jiménez-Guerrero JA. Mercury, lead and cadmium levels in the urine of 170 Spanish adults: a pilot human biomonitoring study. Int J Hyg Environ Health. 2012;215(2):191-5.

86. Castaño A, Sánchez-Rodríguez JE, Cañas A, Esteban M, Navarro C, Rodríguez-García AC, Arribas M, Díaz G, Jiménez-Guerrero JA. 2012. Mercury, lead and cadmium levels in the urine of 170 Spanish adults: a pilot human biomonitoring study. Int J Hyg Environ Health. 215(2):191-5.

87. Castro-Giner F, Künzli N, Jacquemin B, Forsberg B, de Cid R, Sunyer J, Jarvis D, Briggs D, Vienneau D, Norback D, González JR, Guerra S, Janson C, Antó JM, Wjst M, Heinrich

J, Estivill X, Kogevinas M. 2009. Traffic-related air pollution, oxidative stress genes, and asthma (ECHRS). Environ Health Perspect. 117(12):1919-24.

88. Cauchi S, Stücker I, Solas C, Laurent-Puig P, Cénée S, Hémon D, Jacquet M, Kremers P, Beaune P, Massaad-Massade L. 2001. Polymorphisms of human aryl hydrocarbon receptor (AhR) gene in a French population: relationship with CYP1A1 inducibility and lung cancer. Carcinogenesis 22: 1819–1824.
89. Cempel M, Nikel G. 2006. Nickel: A Review of Its Sources and Environmental Toxicology. Contact Dermatitis 15(3):375–382.
90. Cerná M, Krsková A, Cejchanová M, Spěváčková V. Human biomonitoring in the Czech Republic: an overview. Int J Hyg Environ Health. 2012;215(2):109-19.
91. Cesaroni G, Boogaard H, Jonkers S, Porta D, Badaloni C, Cattani G, et al. 2012. Health benefits of traffic-related air pollution reduction in different socioeconomic groups: the effect of low-emission zoning in Rome. Occup Environ Med 69: 133-139.
92. Chappell S, Daly L, Morgan K, Guetta-Baranes T, Roca J, Rabinovich R, Lotya J, Millar a B, Donnelly SC, Keatings V, MacNee W, Stolk J, Hiemstra PS, Miniati M, Monti S, O'Connor CM, Kalsheker N. 2008. Genetic variants of microsomal epoxide hydrolase and glutamate-cysteine ligase in COPD. The European Respiratory Journal 32(4):931–937.
93. Characterization of human lung microsomal cytochrome P-450 1A1 and its role in the oxidation of chemical carcinogens. Molecular Pharmacology 41: 856–864.
94. Chartier J, Pinard M. 2005. OrageneTM and REPLI-g® whole genome amplification. DNA Genotek, Inc.; Ottawa.
95. Chauhan AJ, Inskip HM, Linaker CH, Smith S, Schreiber J, Johnston SL, Holgate ST. 2003. Personal exposure to nitrogen dioxide (NO2) and the severity of virus-induced asthma in children. Lancet 361:1939–1944.
96. Chay KY, Greenstone M. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. Q J Econ. 2003;118:1121–1167.
97. Chen B, Hu Y, Jin T, Lu D, Shao M, Zheng L, Wang Q, Shen Y, Liu H, Liu Y, Zhou Y. 2007a. The influence of metabolic gene polymorphisms on urinary 1-hydroxypyrene concentrations in Chinese coke oven workers. Science of the Total Environment 381: 38–46.
98. Chen C, Arjomandi M, Tager IB, Holland N, Balmes JR. 2007. Effects of antioxidant enzyme polymorphisms on ozone-induced lung function changes. The European Respiratory Journal 30(4):677–683.

99. Chen D, Tian T, Wang H, Liu H, Hu Z, Wang Y, Liu Y, Ma H, Fan W, Miao R, Sun W, Qian J, Jin L, Wei Q, Shen H, Huang W, Lu D. 2009. Association of human aryl hydrocarbon receptor gene polymorphisms with risk of lung cancer among cigarette smokers in a Chinese population. Pharmacogenetics and Genomics 19(1):25–34.
100. Chen X, Abdulhamid I, Woodcroft K. 2011. Maternal smoking during pregnancy, polymorphic CYP1A1 and GSTM1, and lung-function measures in urban family children. Environmental Research 111(8):1215–1221.
101. Cheng SL, Yu CJ, Chen CJ, Yang PC. 2004. Genetic polymorphism of epoxide hydrolase and glutathione S-transferase in COPD. Eur Respir J. 23:818–824.
102. Chiang HC, Huang YK, Chen PF, Chang CC, Wang CJ, Lin P, Lee HL. 2012. 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone is correlated with 8-hydroxy-2'-deoxyguanosine in humans after exposure to environmental tobacco smoke. Sci Total Environ. 414:134-9.
103. Chiba M, Masironi R. 1992. Toxic and trace elements in tobacco and tobacco smoke. Bulletin of the World Health Organization 70: 269–275.
104. Cho HY, Kleeberger SR. 2010. Nrf2 protects against airway disorders. Toxicol Appl Pharmacol. 244(1):43-56.
105. Cho HY, Reddy SP, Kleeberger SR. 2006. Nrf2 defends the lung from oxidative stress. Antioxid Redox Signal. 8(1-2):76-87.
106. Cho H-Y, Reddy SPM, Yamamoto M, Kleeberger SR. 2004. The transcription factor NRF2 protects against pulmonary fibrosis. FASEB J 18(11):1258–1260.
107. Choi MK, Kang MH, Kim MH. 2009. The analysis of copper, selenium, and molybdenum contents in frequently consumed foods and an estimation of their daily intake in korean adults. Biol Trace Elem Res. 128(2):104-17.
108. Chuang CY, Lee CC, Chang YK, Sung FC. 2003. Oxidative DNA damage estimated by urinary 8-hydroxydeoxyguanosine: influence of taxi driving, smoking and areca chewing. Chemosphere. 52(7):1163-71.
109. Ciccone G, Forastiere F, Agabiti N, Biggeri A, Bisanti L, Chellini E, Corbo G, Orco VD, Dalmasso P, Volante TF, Galassi C, Piver S, Renzoni E, Rusconi F. 1998. Road traffic and adverse respiratory effects in children. SIDRIA Collaborative Group. Occupational and Environmental Medicine 5: 771–778.
110. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: An intervention study. Lancet. 2002;360:1210–1214.

111. Clark NA, Demers PA, Karr CJ, Koehoorn M, Lencar C, Tamburic L, Brauer M. Effect of early life exposure to air pollution on development of childhood asthma. Environ Health Perspect 2010;118:284–290.
112. Cohen EA, Linda H, Burke JM, Mccurdy TR, Berry MR, Rigas ML, Zartarian VG, Freema NCG. 2000. Children's exposure assessment: a review of factors influencing Children's exposure, and the data available to characterize and assess that exposure. Environmental Health Perspectives 108: 475–486.
113. Collins FS. 2004. What we do and don ' t know about " race ", " ethnicity ", genetics and health at the dawn of the genome era. Nature Genetics 36(11):13–15.
114. Comhair S, Erzurum SC. 2002. Antioxidant responses to oxidant-mediated lung diseases. American Journal of Physiology Lung Cellular and Molecular Physiology 283(2):L246–L255.
115. Commodore AA, Zhang JJ, Chang Y, Hartinger SM, Lanata CF, Mäusezahl D, Gil AI, Hall DB, Aguilar-Villalobos M, Vena JE, Wang JS, Naeher LP. 2013. Concentrations of urinary 8-hydroxy-2'-deoxyguanosine and 8-isoprostane in women exposed to woodsmoke in a cookstove intervention study in San Marcos, Peru. Environ Int. 60:112-22.
116. Cooke MS, Lunec J, Evans MD. 2002. Progress in the analysis of urinary oxidative DNA damage. Free Radical Biology & Medicine 33: 1601–1614.
117. Cooper RG, Harrison AP. 2009. The exposure to and health effects of antimony. Indian Journal of Occupational and Environmental Medicine 13: 3–10.
118. Corbett J. 1995. Accidental poisoning with iron supplements. MCN, American Journal of Maternal Child Nursing 20: 234.
119. Corradi M, Alinovi R, Goldoni M, Vettori M, Folesani G, Mozzoni P, Cavazzini S, Bergamaschi E, Rossi L, Mutti A. 2002. Biomarkers of oxidative stress after controlled human exposure to ozone. Toxicol Lett. 134:219–225.
120. Correia-Deur JE, Claudio L, Imazawa AT, Eluf-Neto J. Variations in peak expiratory flow measurements associated to air pollution and allergic sensitization in children in Sao Paulo, Brazil. Am J Ind Med. 2012 Dec;55(12):1087-98.
121. Costa DL, Dreher KL. 1997. Bioavailable transition metals in particulate matter mediate cardiopulmonary injury in healthy and compromised animal models. Environmental Health Perspectives 105: 1053–1060.
122. Councell TB, Duckenfield KU, Landa ER, Callender E. 2004. Tire-wear particles as a source of zinc to the environment. Environmental Science & Technology 38: 4206–4214.

123. Cunningham JM, Sellers TA, Schildkraut JM, Zachary S, Vierkant RA, Kelemen LE, Gadre M, Catherine M, Huang Y, Meyer JG, Pankratz VS, Goode EL. 2008. Performance of amplified DNA in an Illumina GoldenGate BeadArray assay. Cancer Epidemiology, Biomarkers & Prevention 17(7):1781–1789.

124. Curjuric I, Imboden M, Nadif R, Kumar A, Schindler C, Haun M, Kronenberg F, Künzli N, Phuleria H, Postma DS, Russi EW, Rochat T, Demenais F, Probst-Hensch NM. 2012. Different genes interact with particulate matter and tobacco smoke exposure in affecting lung function decline in the general population. PLoS ONE 7(7):e40175.

125. Dahl M, Bowler RP, Juul K, Crapo JD, Levy S, Nordestgaard BG. 2008. Superoxide dismutase 3 polymorphism associated with reduced lung function in two large populations. American Journal of Respiratory and Critical Care Medicine 178(9):906–912.

126. Dales R, Wheeler AJ, Mahmud M, Frescura AM, Liu L. 2009. The influence of neighborhood roadways on respiratory symptoms among elementary schoolchildren. J Occup Environ Med 51:654-660.

127. David GL, Romieu I, Sienra-Monge JJ, Collins WJ, Ramirez-Aguilar M, Rio-Navarro BE del, Reyes-Ruiz NI, Morris RW, Marzec JM, London SJ. 2003. Nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase and glutathione S-transferase M1 polymorphisms and childhood asthma. American Journal of Respiratory and Critical Care Medicine 168(10):1199–1204.

128. de Onis M, Onyango A, Borghi E, Siyam A, Blossner M, Lutter C. Worldwide implementation of the WHO child growth standards. Public Health Nutr 2012;12:1–8.

129. Delfino RJ, Chang J, Wu J, Ren C, Tjoa T, Nickerson B, et al. 2009. Repeated hospital encounters for asthma in children and exposure to traffic-related air pollution near the home. Ann Allergy Asthma Immunol 102:138-144.

130. Denison MS, Nagy SR. 2003. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. Annual Review of Pharmacology and Toxicology 43:309–334.

131. Department for Communities and Local Government. 2011. The English Indices of Deprivation 2010 Available: https://www.gov.uk/government/publications/english-indices-of-deprivation-2010 [accessed 13 April 2011].

132. Department of Health. 2012a. Health Profile 2012 Tower Hamlets. Available: http://www.apho.org.uk/default.aspx?QN=HP_METADATA&AreaID=50293.

133. Department of Health. 2012b. Health Profile 2012 Hackney. Available: http://www.apho.org.uk/default.aspx?QN=HP_METADATA&AreaID=50275.

134. Dillon MA, Harris B, Hernandez ML, Zou B, Reed W, Bromberg PA, Devlin RB, Diaz-Sanchez D, Kleeberger S, Zhou H, Lay JC, Alexis NE, Peden DB. 2011. Enhancement of systemic and sputum granulocyte response to inhaled endotoxin in people with the GSTM1 null genotype. Occup Environ Med. 68:783–785.
135. Dockery DW, Cunningham J, Damokosh a I, Neas LM, Spengler JD, Koutrakis P, Ware JH, Raizenne M, Speizer FE. 1996. Health effects of acid aerosols on North American children: respiratory symptoms. Environmental Health Perspectives 104: 500–505.
136. Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG, Speizer FE. 1993. An association between air pollution and mortality in six U.S. cities. The New England Journal of Medicine 329(24):1753–1759.
137. Dockery DW, Speizer FE, Stram DO, Ware JH, Spengler JD, Ferris BG Jr. Effects of inhalable particles on respiratory health of children. Am Rev Respir Dis 1989;139:587–94.
138. Downs SH, Schindler C, Liu LJ, Keidel D, Bayer-Oglesby L, Brutsche MH, Gerbase MW, Keller R, Künzli N, Leuenberger P, Probst-Hensch NM, Tschopp JM, Zellweger JP, Rochat T, Schwartz J, Ackermann-Liebrich U. 2007. Reduced exposure to PM10 and attenuated age-related decline in lung function. N Engl J Med 357:2338–2347.
139. Drake K, Galanter JM, Burchard EG. 2008. Race, ethnicity and social class and the complex etiologies of asthma. Pharmacogenomics 9(4):453–462.
140. Duhme H, Weiland SK, Keil U, Kraemer B, Schmid M, Stender M, Chambless L. 1996. The association between self-reported symptoms of asthma and allergic rhinitis and self-reported traffic density on street of residence in adolescents. Epidemiology 7: 578–82.
141. Dundas I, Chan EY, Bridge PD, McKenzie SA. 2005. Diagnostic accuracy of bronchodilator responsiveness in wheezy children. Thorax 60:13-16.
142. Dunnill MS. 1962. Postnatal Growth of the Lung. Thorax 17(4):329–333.
143. Eenhuizen E, Gehring U, Wijga AH, Smit HA, Fischer PH, Brauer M, Koppelman GH, Kerkhof M, de Jongste JC, Brunekreef B, Hoek G. Traffic related air pollution is related to interrupter resistance in four-year old children. Eur Respir J. 2013. 41(6):1257-63.
144. Ekelund U, Yngve A, Brage S, Westerterp K, Sjöström M. 2004. Body movement and physical activity energy expenditure in children and adolescents: how to adjust for differences in body size and age. The American Journal of Clinical Nutrition 79: 851–6.
145. Environmental Protection Agency Centers for Children's Environmental Health and Disease Prevention Research. Environmental Health Perspectives 113(10):1414–1418.
146. Environmental Research Group KCL. London Air. Available: http://www.londonair.org.uk [accessed 10/07/2013].

147. Ewing P, Otczyk DC, Occhipinti S, Kyd JM, Gleeson M, Cripps AW. 2010. Developmental profiles of mucosal immunity in pre-school children. Clin Dev Immunol. 2010:196785.

148. Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–87.

149. Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7: 574–578.

150. Fan J-B, Hu S, Craumer W, Barker D. 2005. BeadArray-based solutions for enabling the promise of pharmacogenomics. Biotechniques 39(4):S583–S588.

151. Fan R, Wang D, Mao C, Ou S, Lian Z, Huang S, Lin Q, Ding R, She J. 2012. Preliminary study of children's exposure to PAHs and its association with 8-hydroxy-2'-deoxyguanosine in Guangzhou, China. Environ Int. 42:53-8.

152. Fanucchi M V, Buckpitt AR, Murphy ME, Plopper CG. 1997. Naphthalene cytotoxicity of differentiating Clara cells in neonatal mice. Toxicology and Applied Pharmacology 144(1):96–104.

153. Fanucchi M V, Day KC, Clay CC, Plopper CG. 2004. Increased vulnerability of neonatal rats and mice to 1-nitronaphthalene-induced pulmonary injury. Toxicology and Applied Pharmacology 201(1):53–65.

154. Farber HJ, Knowles SB, Brown NL, Caine L, Luna V, Qian Y, Lavori P, Wilson SR. 2008. Secondhand tobacco smoke in children with asthma: sources of and parental perceptions about exposure in children and parental readiness to change. Chest 133(6):1367–74.

155. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souëf P, Danoy P, Baltic S, Nyholt DR, Jenkins M, Hayden C, Willemsen G, Ang W, Kuokkanen M, Beilby J, Cheah F, de Geus EJ, Ramasamy A, Vedantam S, Salomaa V, Madden PA, Heath AC, Hopper JL, Visscher PM, Musk B, Leeder SR, Jarvelin MR, Pennell C, Boomsma DI, Hirschhorn JN, Walters H, Martin NG, James A, Jones G, Abramson MJ, Robertson CF, Dharmage SC, Brown MA, Montgomery GW, Thompson PJ; Australian Asthma Genetics Consortium. 2011. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet 201: 1006–1014.

156. Figueroa JD, Malats N, García-Closas M, Real FX, Silverman D, Kogevinas M, Chanock S, Welch R, Dosemeci M, Lan Q, Tardón A, Serra C, Carrato A, García-Closas R,

Castaño-Vinyals G, Rothman N. 2008. Bladder cancer risk and genetic variation in AKR1C3 and other metabolizing genes. Carcinogenesis 29: 1955–1962.

157. Finkelstein MM, Jerrett M, Sears MR. 2004. Traffic air pollution and mortality rate advancement periods. American Journal of Epidemiology 160(2):173–177.

158. Figarska SM, Vonk JM, Boezen HM. 2014. NFE2L2 polymorphisms, mortality, and metabolism in the general population. Physiol Genomics. 15;46(12):411-7.

159. Folz RJ, Abushamaa AM, Suliman HB. 1999. Extracellular superoxide dismutase in the airways of transgenic mice reduces inflammation and attenuates lung toxicity following hyperoxia. Journal of Clinical Investigation 103: 1055–1066.

160. Freire C, Abril A, Fernández MF, Ramos R, Estarlich M, Manrique A, Aguirre A, Ibarluzea J, Olea N. 2009. Urinary 1-hydroxypyrene and PAH exposure in 4-year-old Spanish children. Science of the Total Environment 407(5):1562–1569.

161. Fritz GJ, Herbarth O. 2001. Pulmonary function and urban air pollution in preschool children. Int J Hyg Environ Health 203:235–244.

162. Fryer AA, Bianco A, Hepple M, Jones PW, Strange RC, Spiteri MA. 2000. Polymorphism at the glutathione S-transferase GSTP1 locus. A new marker for bronchial hyperresponsiveness and asthma. American Journal of Respiratory and Critical Care Medicine 161: 1437–1442.

163. Fryer AA, Hume R, Strange RC. 1986. The development of glutathione S-transferase and glutathione peroxidase activities in human lung. Biochim Biophys Acta. 883:448–453

164. Fujii-Kuriyama Y, Mimura J. 2005. Molecular mechanisms of AhR functions in the regulation of cytochrome P450 genes. Biochemical and Biophysical Research Communications 338(1):311–317.

165. Gabriele C, Asgarali R, Jaddoe VW, Hofman A, Moll HA, de Jongste JC. Smoke exposure, airway symptoms and exhaled nitric oxide in infants: the Generation R study. Eur Respir J 2008;32:307–313. 19.

166. Gagliardi L, Rusconi F. 1997. Respiratory rate and body mass in the first three years of life. Archives of Disease in Childhood 76(2):151–154.

167. Ganguly K, Depner M, Fattman C, Bein K, Oury TD, Wesselkamper SC, Borchers MT, Schreiber M, Gao F, Mutius E von, Kabesch M, Leikauf GD, Schulz H. 2009. Superoxide dismutase 3, extracellular (SOD3) variants and lung function. Physiological Genomics 37(3):260–267.

168. Gao Y, Chan EY, Li LP, He QQ, Wong TW. 2013. Chronic effects of ambient air pollution on lung function among Chinese children. Arch Dis Child. 98(2):128-35.

169. Gauderman WJ, Avol E, Gilliland F, Vora H, Thomas D, Berhane K, et al. 2004. The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med 351:1057-1067.

170. Gauderman WJ, Avol E, Lurmann F, Kuenzli N, Gilliland F, Peters J, McConnell R. 2005. Childhood Asthma and Exposure to Traffic and Nitrogen Dioxide. Epidemiology 16(6):737–743.

171. Gauderman WJ, Gilliland GF, Vora H, Avol E, Stram D, McConnell R, Thomas D, Lurmann F, Margolis HG, Rappaport EB, Berhane K, Peters JM. 2002. Association between air pollution and lung function growth in southern California children: results from a second cohort. Am J Respir Crit Care Med. 166(1):76-84.

172. Gauderman WJ, McConnell R, Gilliland F, London S, Thomas D, Avol E, Vora H, Berhane K, Rappaport EB, Lurmann F, Margolis HG, Peters J. 2000. Association between air pollution and lung function growth in southern California children. American Journal of Respiratory and Critical Care Medicine 162: 1383–1390.

173. Gauderman WJ, Vora H, McConnell R, Berhane K, Gilliland F, Thomas D, et al. 2007. Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study. Lancet 369:571-577.

174. Gauvin S, Amro S, Zmirou D, Le Moullec Y, Sahraoui F, Pin I, Neukirch F, Momas I, Lauvergne N, Poilve M-P, Chiron M. 2001. Road traffic, NO_2 exposure and respiratory function among children (VESTA study). Int J Vehicle Des 27:251–261.

175. Gehring U, Cyrys J, Sedlmeir G, Brunekreef B, Bellander T, Fischer P, Bauer CP, Reinhardt D, Wichmann HE, Heinrich J. Traffic-related air pollution and respiratory health during the first 2 yrs of life. 2002. The European Respiratory Journal 19: 690–698.

176. Gehring U, Gruzieva O, Agius RM, Beelen R, Custovic A, Cyrys J, Eeftens M, Flexeder C, Fuertes E, Heinrich J, Hoffmann B, de Jongste JC, Kerkhof M, Klümper C, Korek M, Mölter A, Schultz ES, Simpson A, Sugiri D, Svartengren M, von Berg A, Wijga AH, Pershagen G, Brunekreef B. Air pollution exposure and lung function in children: the ESCAPE project. 2013. Environmental Health Perspectives. 121(11-12):1357-64

177. Gerde P, Muggenburg BA, Lundborg M, Tesfaigzi Y, Dahl AR. 2001. Respiratory epithelial penetration and clearance of particle-borne benzo[a]pyrene. Research Report Health Effects Institute 5–25; discussion 27–32.

178. Gerlofs-Nijland ME, Dormans JAMA, Bloemen HJT, Leseman DLAC, John A, Boere F, Kelly FJ, Mudway IS, Jimenez AA, Donaldson K, Guastadisegni C, Janssen NAH, Brunekreef B, Sandström T, Bree L Van, Cassee FR. 2007. Toxicity of coarse and fine

particulate matter from sites with contrasting traffic profiles. Inhalation Toxicology 19: 1055–1069.

179. Ghio AJ, Hilborn ED, Stonehuerner JG, Dailey LA, Carter JD, Richards JH, Crissman KM, Foronjy RF, Uyeminami DL, Pinkerton KE. 2008. Particulate matter in cigarette smoke alters iron homeostasis to produce a biological effect. Am J Respir Crit Care Med. 178(11):1130-8.

180. Ghio AJ, Suliman HB, Carter JD, Abushamaa AM, Folz RJ. 2002. Overexpression of extracellular superoxide dismutase decreases lung injury after exposure to oil fly ash. American Journal of Physiology Lung Cellular and Molecular Physiology 283(1):L211–L218.

181. Gibson RS. 2012. A historical review of progress in the assessment of dietary zinc intake as an indicator of population zinc status. Adv Nutr. 2012. 3(6):772-82.

182. Gilliland FD, Gauderman WJ, Vora H, Rappaport E, Dubeau L. 2002. Effects of glutathione-S-transferase M1, T1, and P1 on childhood lung function growth. Am J Respir Crit Care Med. 166:710–716.

183. Gilliland FD, Li YF, Gong H Jr, Diaz-Sanchez D. 2006. Glutathione s-transferases M1 and P1 prevent aggravation of allergic responses by secondhand smoke. Am J Respir Crit Care Med.174 (12):1335– 1341

184. Gilliland FD, Li Y-F, Saxon A, Diaz-Sanchez D. 2004. Effect of glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement of allergic responses: randomised, placebo-controlled crossover study. Lancet 363: 119–125.

185. Gilmour MI, Jaakkola MS, London SJ, Nel AE, Rogers C. 2006. How Exposure to Environmental Tobacco Smoke, Outdoor Air Pollutants, and Increased Pollen Burdens Influences the Incidence of Asthma. Environmental Health Perspectives 114(4):627–633.

186. Ginsberg GL, Foos BP, Firestone MP. 2005. Review and analysis of inhalation dosimetry methods for application to children's risk assessment. Journal Of Toxicology And Environmental Health Part A 68(8):573–615.

187. Glinianaia S V., Rankin J, Bell R, Pless-Mulloli T, Howel D. 2004. Does Particulate Air Pollution Contribute to Infant Death? A Systematic Review. Environmental Health Perspectives 112(14):1365–1370.

188. Godri KJ, Harrison RM, Evans T, Baker T, Dunster C, Mudway IS, Kelly FJ. 2011. Increased oxidative burden associated with traffic component of ambient particulate matter at roadside and urban background schools sites in London. PLoS One. 6(7):e21961.

189. Götschi T, Heinrich J, Sunyer J, Künzli N. Long-term effects of ambient air pollution on lung function: a review. Epidemiology. 2008. 19(5):690-701.

190. Greater London Authority (GLA). The London Atmospheric Emissions Inventory 2008 http://data.london.gov.uk/laei-2008

191. Green DC, Fuller GW, Baker T. Development and validation of the volatile correction model for PM10 - An empirical method for adjusting TEOM measurements for their loss of volatile particulate matter. 2009. Atmos Environ. 43(13):2132-2141. NHS Information Centre for Health and Social Care. 2011. Compendium of Population Health Indicators. Available: https://indicators.ic.nhs.uk/webview/.

192. Grigg J. Air pollution and children's respiratory health - gaps in the global evidence. 2011. Clin Exp Allergy. 41(8):1072-5.

193. Gu A, Ji G, Jiang T, Lu A, You Y, Liu N, Luo C, Yan W, Zhao P. 2012. Contributions of aryl hydrocarbon receptor genetic variants to the risk of glioma and PAH-DNA adducts. Toxicological Sciences 128(2):357–364.

194. Guo F, Dey DK, Holsinger KE. 2009. A Bayesian hierarchical model for analysis of SNP diversity in multilocus, multipopulation samples. Journal of the American Statistical Association 104(485):142–154.

195. Halliwell B. 1991. Reactive oxygen species in living systems: source, biochemistry, and role in human disease. The American Journal of Medicine 91: 14S–22S.

196. Han YY, Donovan M, Sung FC. 2010. Increased urinary 8-hydroxy-2'-deoxyguanosine excretion in long-distance bus drivers in Taiwan. Chemosphere. 79(9):942-8.

197. Hanene C, Jihene L, Jamel A, Kamel H, Agnes H. 2007. Association of GST genes polymorphisms with asthma in Tunisian children. Mediators Inflamm. 2007:19564.

198. Hankinson O: The aryl hydrocarbon receptor complex. 1995. Ann Rev Pharmacol Toxicol. 35:307-340.

199. Hansen AM, Raaschou-Nielsen O, Knudsen LE. 2005. Urinary 1-hydroxypyrene in children living in city and rural residences in Denmark. Science of the Total Environment 347(1-3):98–105.

200. Hansen TVO, Simonsen MK, Nielsen FC, Hundrup YA. 2007. Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA. Cancer Epidemiology, Biomarkers & Prevention 16(10):2072–2076.

201. Harper PA, Wong JMY, Lam MSM, Okey AB. 2002. Polymorphisms in the human AH receptor. Chemicobiological Interactions 141: 161–187.

202. Harrison RM, Yin J. 2000. Particulate matter in the atmosphere: which particle properties are important for its effects on health? Science of the Total Environment 249: 85–101.

203. Hassett C, Aicher L, Sidhu JS, Omiecinski CJ. 1994. Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants. Human Molecular Genetics 3: 421–428.

204. Hayashi S, Watanabe J, Nakachi K, Kawajiri K. 1991. Genetic linkage of lung cancer-associated MspI polymorphisms with amino acid replacement in the heme binding region of the human cytochrome P450IA1 gene. Journal of Biochemistry 110: 407–411.

205. Hayes JD, Strange RC. 2000. Glutathione S-transferase polymorphisms and their biological consequences. Pharmacology 61(3):154–166.

206. He QQ, Wong TW, Du L, Jiang ZQ, Gao Y, Qiu H, Liu WJ, Wu JG, Wong A, Yu TS. 2010. Effects of ambient air pollution on lung function growth in Chinese schoolchildren. Respir Med. 104:1512–20.

207. Health Effects Institute. Traffic-related air pollution: a critical review of the literature on emissions, exposure and health effects. Boston, MA: Health Effects Institute. 2010. HEI Special Report 17. pp. 1–115.

208. Hedley AJ, Wong CM, Thach TQ, Ma S, Lam TH, Anderson HR. 2002. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: An intervention study. Lancet. 360:1646–1652.

209. HEI Panel on the Health Effects of Traffic-Related Air Pollution. 2010. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. Boston:Health Effects Institute.

210. Heinrich J, Hoelscher B, Wjst M, Ritz B, Cyrys J, Wichmann H. 1999. Respiratory diseases and allergies in two polluted areas in East Germany. Environ Health Perspect. 107(1):53-62.

211. Henderson a J, Newson RB, Rose-Zerilli M, Ring SM, Holloway JW, Shaheen SO. 2010. Maternal Nrf2 and gluthathione-S-transferase polymorphisms do not modify associations of prenatal tobacco smoke exposure with asthma and lung function in school-aged children. Thorax 65(10):897–902.

212. Henderson FW, Reid HF, Morris R, Wang OL, Hu PC, Helms RW, et al. 1989. Home air nicotine levels and urinary cotinine excretion in preschool children. Am Rev Respir Dis 140:197-201.

213. Higgins MW, Keller JB. 1970. Predictors of mortality in the adult population of Tecumseh; Respiratory symptoms, chronic respiratory disease, and ventilatory lung function. Arch Environ Health 21:418–424.

214. Higashi T, Shibayama Y, Fuji M, Shimada K. 2008. Liquid chromatography-tandem mass spectrometric method for the determination of salivary 25-hydroxyvitamin D3: a noninvasive tool for the assessment of vitamin D status. Anal Bioanal Chem. 391(1):229-38.

215. High-throughput SNP genotyping on universal bead arrays. Mutation Research 573(1-2):70–82.

216. Hirsch T, Weiland SK, von Mutius E, Safeca AF, Gräfe H, Csaplovics E, Duhme H, Keil U, Leupold W. 1999. Inner city air pollution and respiratory health and atopy in children. Eur Respir J 14:669–677.

217. Hoek G, Brunekreef B, Goldbohm S, Fischer P, Brandt P a van den. 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. Lancet 360(9341):1203–1209.

218. Hoek G, Pattenden S, Willers S, Antova T, Fabianova E, Braun-Fahrlander C, et al. 2012. PM10, and children's respiratory symptoms and lung function in the PATY study. Eur Respir J 40:538-547.

219. Hoffmann B, Moebus S, Möhlenkamp S, Stang A, Lehmann N, Dragano N, Schmermund A, Memmesheimer M, Mann K, Erbel R, Jöckel KH, Heinz Nixdorf Recall Study Investigative Group. 2007. Residential exposure to traffic is associated with coronary atherosclerosis. Circulation. 116(5):489-96.

220. Hoffmann B, Moebus S, Stang A, Beck EM, Dragano N, Möhlenkamp S, Schmermund A, Memmesheimer M, Mann K, Erbel R, Jöckel KH; Heinz Nixdorf RECALL Study Investigative Group. 2006. Residence close to high traffic and prevalence of coronary heart disease. Eur Heart J. 27(22):2696-702.

221. Hoidal JR. 2001. Perspective Reactive Oxygen Species and Cell Signaling. American Journal of Respiratory Cell and Molecular Biology 25: 661–663.

222. Hole DJ, Watt GC, Davey-Smith G, Hart CL, Gillis CR, Hawthorne VM. 1996. Impaired lung function and mortality risk in men and women: Findings from the Renfrew and Paisley prospective population study. Br Med J (Clin Res Ed) 313:711–715.

223. Holgate ST, Devlin RB, Wilson SJ, Frew AJ. 2003. Health effects of acute exposure to air pollution. Part II: Healthy subjects exposed to concentrated ambient particles. Res Rep Health Eff Inst. (112):31-50

224. Holgate ST, Sandström T, Frew AJ, Stenfors N, Nördenhall C, Salvi S, Blomberg A, Helleday R, Söderberg M. 2003. Health effects of acute exposure to air pollution. Part I: Healthy and asthmatic subjects exposed to diesel exhaust. Res Rep Health Eff Inst. (112):1-30

225. Holguin F, Flores S, Ross Z, Cortez M, Molina M, Molina L, Rincon C, Jerrett M, Berhane K, Granados A, Romieu I. 2007. Traffic-related exposures, airway function, inflammation, and respiratory symptoms in children. Am J Respir Crit Care Med 176:1236–1242.

226. Holsinger KE, Lewis PO. 2002. A Bayesian approach to inferring population structure from docimant markers structure from dominant markers. Molecular Ecology 11: 1157–1164.

227. Holsinger KE, Weir BS. 2009. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nature Reviews Genetics 10(9):639–650.

228. Hong Y-C, Hwang S-S, Kim JH, Lee K-H, Lee H-J, Lee K-H, Yu S-D, Kim D-S. 2007. Metals in Particulate Pollutants Affect Peak Expiratory Flow of Schoolchildren. Environmental Health Perspectives 115: 430–434.

229. Hong YC, Leem JH, Lee KH, Park DH, Jang JY, Kim ST, Ha EH. 2005. Exposure to air pollution and pulmonary function in university students. Int Arch Occup Environ Health 78:132–138.

230. Hope BK. 1997. An assessment of the global impact of anthropogenic vanadium. Biogeochemistry 37: 1–13.

231. Horak FJr, Studnicka M, Gartner C, et al. 2002. Particulate matter and lung function growth in children: a 3-yr follow-up study in Austrian schoolchildren. Eur Respir J. 19:838–845.

232. Hou SM, Ryberg D, Falt S, Deverill A, Tefre T, Borresen AL, Haugen A, Lambert B. 2000. GSTM1 and NAT2 polymorphisms in operable and non-operable lung cancer patients. Carcinogenesis. 21:49–54.

233. Huang H, Helzlsouer KJ, Appel LJ. 2000. The effects of vitamin C and vitamin E on oxidative DNA damage: results from a randomized controlled trial. Cancer Epidemiology, Biomarkers & Prevention 9: 647–652.

234. Huang HB, Chen GW, Wang CJ, Lin YY, Liou SH, Lai CH, Wang SL. 2013. Exposure to heavy metals and polycyclic aromatic hydrocarbons and DNA damage in Taiwanese traffic conductors. Cancer Epidemiol Biomarkers Prev. 22(1):102-8.

235. Huang HB, Lai CH, Chen GW, Lin YY, Jaakkola JJ, Liou SH, Wang SL. 2012. Traffic-related air pollution and DNA damage: a longitudinal study in Taiwanese traffic conductors. PLoS One. 7(5):e37412.

236. Huang HY, Appel LJ, Croft KD, Miller ER 3rd, Mori TA, Puddey IB. 2002. Effects of vitamin C and vitamin E on in vivo lipid peroxidation: results of a randomized controlled trial. Am J Clin Nutr. 76(3):549-55.

237. Huang HY, Helzlsouer KJ, Appel LJ. 2000. The effects of vitamin C and vitamin E on oxidative DNA damage: results from a randomized controlled trial. Cancer Epidemiol Biomarkers Prev. 9(7):647-52.

238. Huang RS, Chen P, Wisel S, Duan S, Zhang W, Cook EH, Das S, Cox NJ, Dolan ME. 2009. Population-specific GSTM1 copy number variation. Human Molecular Genetics 18: 366–372.

239. Huang W, Wang G, Lu SE, Kipen H, Wang Y, Hu M, Lin W, Rich D, Ohman-Strickland P, Diehl SR, Zhu P, Tong J, Gong J, Zhu T, Zhang J. 2012. Inflammatory and oxidative stress responses of healthy young adults to changes in air quality during the Beijing Olympics. Am J Respir Crit Care Med. 186(11):1150-9.

240. Hurrell RF. 1997. Preventing iron deficiency through food fortification. Nutrition Reviews 55: 210–222.

241. Hutcheon JA, Chiolero A, Hanley JA. 2010. Random measurement error and regression dilution bias. BMJ. 340:c2289.

242. Iavicoli I, Leso V, Fontana L, Bergamaschi A. 2011. Toxicological effects of titanium dioxide nanoparticles: a review of in vitro mammalian studies. European Review for Medical and Pharmacological Sciences 15: 481–508.

243. Iba MM, Shin M, Caccavale RJ: Cytochromes P4501 (CYP1): Catalytic activities and inducibility by diesel exhaust particle extract and benzo[a]pyrene in intact human lung ex vivo. 2010. Toxicology 273:35-44/

244. Igishi T, Hitsuda Y, Kato K, Sako T, Burioka N, Yasuda K, Sano H, Shigeoka Y, Nakanishi H, Shimizu E. 2003. Elevated urinary 8-hydroxydeoxyguanosine, a biomarker of oxidative stress, and lack of association with antioxidant vitamins in chronic obstructive pulmonary disease. Respirology. 8(4):455-60.

245. Iliff A, Lee VA. 1952. Pulse Rate , Respiratory Rate , and Body Temperature of Children between Two Months and Eighteen Years of Age. Child Dev 23: 237–245.

246. Illumina Technical Note. 2006. GoldenGate Assay Workflow.

247. Il'yasova D, Scarbrough P, Spasojevic I. 2012. Urinary biomarkers of oxidative status. Clin Chim Acta. 413(19-20):1446-53.

248. International Agency for Research on Cancer. IARC: Diesel engine exhaust carcinogenic. Lyon, France: International Agency for Research on Cancer; 2012 [accessed 2012 Nov 16]. Press release No. 213. Available from: http://press.iarc.fr/pr213_E.pdf

249. Ip MS, Karlberg EM, Karlberg JP, Luk KD, Leong JC. 2000. Lung function reference values in Chinese children and adolescents in Hong Kong. I. Spirometric values and comparison with other populations. Am J Respir Crit Care Med. 162:424–9.

250. Ising H, Lange-Asschenfeldt H, Lieber GF, Weinhold H, Eilts M. 2003. Respiratory and dermatological diseases in children with long-term exposure to road traffic immissions. Noise Health 5: 41–50.

251. Iskandar A, Andersen ZJ, Bonnelykke K, Ellermann T, Andersen KK, Bisgaard H. 2012. Coarse and fine particles but not ultrafine particles in urban air trigger hospital admission for asthma in children. Thorax 67: 252-257.

252. Islam MS, Ahmed MU, Sayeed MS Bin, Maruf A Al, Mostofa a GM, Hussain SMA, Kabir Y, Daly AK, Hasnat A. 2013. Lung cancer risk in relation to nicotinic acetylcholine receptor, CYP2A6 and CYP1A1 genotypes in the Bangladeshi population. Clinica Chimica Acta 416:11–19.

253. Islam T, Berhane K, McConnell R, Gauderman WJ, Avol E, Peters JM, Gilliland FD. 2009. Glutathione-S-transferase (GST) P1, GSTM1, exercise, ozone and asthma incidence in school children. Thorax 64(3):197–202.

254. Islam T, Breton C, Salam MT, McConnell R, Wenten M, Gauderman WJ, Conti D, Berg D Van Den, Peters JM, Gilliland FD. 2010. Role of inducible nitric oxide synthase in asthma risk and lung function growth during adolescence. Thorax 65(2):139–145.

255. Islam T, McConnell R, Gauderman WJ, Avol E, Peters JM, Gilliland FD. 2008. Ozone, oxidant defense genes, and risk of asthma during adolescence. American Journal of Respiratory and Critical Care Medicine 177(4):388–395.

256. Iyengar GV. 1998. Reevaluation of the trace element content in reference man. Radiat Phys Chem 51:545–60.

257. Jackson BP, Taylor VF, Karagas MR, Punshon T, Cottingham KL. 2012. Arsenic, organic foods, and brown rice syrup. Environ Health Perspect. 120(5):623-6.

258. Jain RB. 2013. Effect of pregnancy on the levels of urinary metals for females aged 17-39 years old: data from National Health and Nutrition Examination Survey 2003-2010. Journal of Toxicology and Environmental Health. Part A 76(2):86–97.

259. Janssen NA, Brunekreef B, van Vliet P, Aarts F, Meliefste K, Harssema H, et al. 2003. The relationship between air pollution from heavy traffic and allergic sensitization, bronchial hyperresponsiveness, and respiratory symptoms in Dutch schoolchildren. Environ Health Perspect 111: 1512-1518.

260. Jarup L. 2003. Hazards of heavy metal contamination. British Medical Bulletin 68(1):167–182.

261. Jedrychowski W, Flak E, Mroz E. 1999. The adverse effect of low levels of ambient air pollutants on lung function growth in preadolescent children. Environ Health Perspect. 107:669–674.

262. Jedrychowski W, Galas A, Pac A, Flak E, Camman D, Rauh V, Perera F. 2005. Prenatal ambient air exposure to polycyclic aromatic hydrocarbons and the occurrence of respiratory symptoms over the first year of life. European Journal of Epidemiology 20: 775–782.

263. Jones LC, Whitlock JP. 2001. Dioxin-inducible transactivation in a chromosomal setting. Analysis of the acidic domain of the Ah receptor. The Journal of Biological Chemistry 276(27):25037–25042.

264. Joshi S, Kotecha S. 2007. Lung growth and development. Early Human Development 83(12):789–94.

265. Joubert BR, Reif DM, Edwards SW, Leiner KA, Hudgens EE, Egeghy P, Gallagher JE, Hubal EC. 2011. Evaluation of genetic susceptibility to childhood allergy and asthma in an African American urban population. BMC Med Genet. 12:25.

266. Juul K, Tybjaerg-Hansen A, Marklund S, Lange P, Nordestgaard BG. 2006. Genetically increased antioxidative protection and decreased chronic obstructive pulmonary disease. American Journal of Respiratory and Critical Care Medicine 173(8):858–864.

267. Jyrkkänen H-K, Kansanen E, Inkala M, Kivelä AM, Hurttila H, Heinonen SE, Goldsteins G, Jauhiainen S, Tiainen S, Makkonen H, Oskolkova O, Afonyushkin T, Koistinaho J, Yamamoto M, Bochkov VN, Ylä-Herttuala S, Levonen A-L. 2008. Nrf2 regulates antioxidant gene expression evoked by oxidized phospholipids in endothelial cells and murine arteries in vivo. Circulation Research 103(1):e1–e9.

268. K Kawajiri, Watanabe J, Eguchi H, Nakachi K, Kiyohara C, Hayashi S. 1995. Polymorphisms of human AhR gene are not involved in lung cancer. Pharmacogenetics 5: 151–158.

269. Kan H, Heiss G, Rose KM, Whitsel E, Lurmann F, London SJ. 2007. Traffic exposure and lung function in adults: The Atherosclerosis Risk in Communities study. Thorax 62:873–879.

270. Kang D, Lee K-M, Park SK, Berndt SI, Peters U, Reding D, Chatterjee N, Welch R, Chanock S, Huang W-Y, Hayes RB. 2007. Functional variant of manganese superoxide dismutase (SOD2 V16A) polymorphism is associated with prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer study. Cancer epidemiology, Biomarkers & Prevention 16(8):1581–1586.

271. Kaur B, Anderson HR, Austin J, Burr M, Harkins LS, Strachan DP, et al. 1998. Prevalence of asthma symptoms, diagnosis, and treatment in 12-14 year old children across Great Britain (international study of asthma and allergies in childhood, ISAAC UK). BMJ 316:118-124.

272. Kawajiri K, Eguchi H, Nakachi K, Sekiya T, Yamamoto M. 1996. Association of CYP1A1 germ line polymorphisms with mutations of the p53 gene in lung cancer. Cancer Research 56: 72–76.

273. Keith LS, Wohlers DW, Moffett DB, Rosemond ZA. 2007. ATSDR evaluation of potential for human exposure to tungsten. Toxicology and Industrial Health 23: 309–345.

274. Kelly F, Anderson HR, Armstrong B, Atkinson R, Barratt B, Beevers S, Derwent D, Green D, Mudway I, Wilkinson P; HEI Health Review Committee. 2011a. The impact of the congestion charging scheme on air quality in London. Part 1. Emissions modeling and analysis of air pollution measurements. Res Rep Health Eff Inst. (155):5-71.

275. Kelly F, Anderson HR, Armstrong B, Atkinson R, Barratt B, Beevers S, Derwent D, Green D, Mudway I, Wilkinson P; HEI Health Review Committee. 2011b. The impact of the congestion charging scheme on air quality in London. Part 2. Analysis of the oxidative potential of particulate matter. Res Rep Health Eff Inst. (155):73-144.

276. Kelly F, Armstrong B, Atkinson R, Anderson HR, Barratt B, Beevers S, Derwent D, Green D, Mudway I, Wilkinson P, HEI Health Review Committee. 2011c. The London Low Emission Zone Baseline Study. Research Report (Health Effects Institute) 163: 3–79.

277. Kelly F, Armstrong B, Atkinson R, Anderson HR, Barratt B, Beevers S, et al. 2011. The London Low Emission Zone Baseline Study. Boston:Health Effects Institute.

278. Kelly FJ. 2004. Dietary antioxidants and environmental stress. Proc Nutr Soc. 63(4):579-85.

279. Kharitonov SA, Barnes PJ. 2004. Effects of corticosteroids on noninvasive biomarkers of inflammation in asthma and chronic obstructive pulmonary disease. Proc Am Thorac Soc. 1(3):191-9.

280. Kidd JR, Friedlaender FR, Speed WC, Pakstis AJ, La Vega FM De, Kidd KK. 2011. Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples. Investigative Genetics 2(1):1.

281. Kidd PM. 1997. Glutathione: Systemic Protectant Against Oxidative and Free Radical Damage. Alternative Medicine Review 2: 155–176.

282. Kim JH, Kim H, Lee KY, Kang J-W, Lee K-H, Park S-Y, Yoon H-I, Jheon SH, Sung SW, Hong Y-C. 2007. Aryl hydrocarbon receptor gene polymorphisms affect lung cancer risk. Lung Cancer 56(1):9–15.

283. Kim SH, Fletcher RA, Zachariah MR. 2005. Understanding the difference in oxidative properties between flame and diesel soot nanoparticles: the role of metals. Environmental Science & Technology 39: 4021–4026.

284. Kimmel C, Collman GW, Fields N, Eskenazi B. 2005. Lessons Learned for the National Children's Study from the National Institute of Environmental Health Sciences/U.S. Environmental Protection Agency Centers for Children's Environmental Health and Disease Prevention Research. Environmental Health Perspectives 113(10):1414–1418.

285. Kirkby J, Bonner R, Lum S, Bates P, Morgan V, Strunk RC, Kirkham F, Sonnappa S, Stocks J. 2013. Interpretation of pediatric lung function: impact of ethnicity. Pediatr Pulmonol. 48(1):20-6.

286. Kiyohara C, Nakanishi Y, Inutsuka S, Takayama K, Hara N, Motohiro A, Tanaka K, Kono S, Hirohata T. 1998. The relationship between CYP1A1 aryl hydrocarbon hydroxylase activity and lung cancer in a Japanese population. Pharmacogenetics 8: 315–323.

287. Kleeberger SR. 2005. Genetic aspects of pulmonary responses to inhaled pollutants. Exp Toxicol Pathol. 57(Suppl 1):147–153.

288. Knaapen AM, Shi T, Borm PJA, Schins RPF. 2002. Soluble metals as well as the insoluble particle fraction are involved in cellular DNA damage induced by particulate matter. Molecular and Cellular Biochemistry 234-235: 317–326.

289. Kosoy R, Nassir R, Tian C, White PA, Butler LLM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF. 2009. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 30:69-78.

290. Kohen R, Nyska A. 2002. Oxidation of biological systems: oxidative stress phenomena, antioxidants, redox reactions, and methods for their quantification. Toxicologic Pathology 30: 620–650.
291. Köhle C, Bock KW. 2007. Coordinate regulation of Phase I and II xenobiotic metabolisms by the Ah receptor and Nrf2. Biochemical Pharmacology 73(12):1853–1862.
292. Kotecha S. 2000. Lung growth for beginners. Paediatric Respiratory Reviews 1(4):308–313.
293. Krämer U, Koch T, Ranft U, Ring J, Behrendt H. 2000. Traffic-related air pollution is associated with atopy in children living in urban areas. Epidemiology 11: 64–70.
294. Krewski D, Burnett R, Jerrett M, Pope CA, Rainham D, Calle E, Thurston G, Thun M. 2005. Mortality and long-term exposure to ambient air pollution: ongoing analyses based on the American Cancer Society cohort. Journal of Toxicology and Environmental Health. Part A 68(13-14):1093–1109.
295. Kreyling WG, Semmler-Behnke M, W Möller. 2004. Dosimetry and toxicology of inhaled ultrafine particles. Journal of Aerosol Medicine 17(2):140–152.
296. Kulkarni N, Pierse N, Rushton L, Grigg J. 2006. Carbon in airway macrophages and lung function in children. N Engl J Med. 355:21–30.
297. Kulkarni N, Pierse N, Rushton L, Grigg J. 2006. Carbon in airway macrophages and lung function in children. N Engl J Med 355:21-30.
298. Künzli N, Jerrett M, Mack WJ, Beckerman B, LaBree L, Gilliland F, Thomas D, Peters J, Hodis HN. 2005. Ambient air pollution and atherosclerosis in Los Angeles. Environ Health Perspect. 113(2):201-6.
299. Kulkarni N, Pierse N, Rushton L, Grigg J. 2006. Carbon in airway macrophages and lung function in children. N Engl J Med. 355(1):21-30.
300. Kwok P-Y. 2001. Methods for genotyping single nucleotide polymorphisms. Annual Review of Genomics and Human Genetics 2: 235–258.
301. Laden F, Schwartz J, Speizer FE, Dockery DW. 2006. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. American Journal of Respiratory and Critical Care Medicine 173(6):667–672.
302. Lai C-H, Liou S-H, Jaakkola JJK, Huang H-B, Su T-Y, Strickland PT. 2012. Exposure to Polycyclic Aromatic Hydrocarbons Associated with Traffic Exhaust: The Increase of Lipid Peroxidation and Reduction of Antioxidant Capacity. Aerosol and Air Quality Research 12(5):941–950.

303. Lai CH, Liou SH, Lin HC, Shih TS, Tsai PJ, Chen JS, Yang T, Jaakkola JJ, Strickland PT. 2005. Exposure to traffic exhausts and oxidative DNA damage. Occup Environ Med. 62(4):216-22.

304. Lakhdar R, Denden S, Mouhamed MH, Chalgoum A, Leban N, Knani J, Lefranc G, Miled A, Ben Chibani J, Khelil AH. 2011. Correlation of EPHX1, GSTP1, GSTM1, and GSTT1 genetic polymorphisms with antioxidative stress markers in chronic obstructive pulmonary disease. Exp Lung Res. 37:195–204.

305. Landi MT, Bertazzi PA, Shields PG, Clark G, Lucier GW, Garte SJ, Cosma G, Caporaso NE. 1994. Association between CYP1A1 genotype, mRNA expression and enzymatic activity in humans. Pharmacogenetics 4: 242–246.

306. Landrigan PJ, Kimmel C a., Correa A, Eskenazi B. 2003. Children's Health and the Environment: Public Health Issues and Challenges for Risk Assessment. Environmental Health Perspectives 112(2):257–265.

307. Langston C, Kida K, Reed M, Thurlbeck WM. 1984. Human lung growth in late gestation and in the neonate. The American Review of Respiratory Disease 129: 607–613.

308. Ledford DK, Burks AW, Ballow M, Cox L, Wood RA, Casale TB. 2012. Statement regarding "the public health benefits of air pollution control." J Allergy Clin Immunol. 130:24.

309. Lee D, Wallis C, Wexler AS, Schelegle ES, VanWinkle LS, Plopper CG, Fanucchi MV, Kumfer B, Kennedy IM, Chan JK. 2010. Small particles disrupt postnatal airway development. J Appl Physiol. 109:1115– 1124.

310. Lee K-M, Ward MH, Han S, Ahn HS, Kang HJ, Choi HS, Shin HY, Koo H-H, Seo J-J, Choi J-E, Ahn Y-O, Kang D. 2009. Paternal smoking, genetic polymorphisms in CYP1A1 and childhood leukemia risk. Leukemia Research 33(2):250–258.

311. Lee MS, Eum KD, Lee K, Kim H, Paek D. 2009. Seasonal and regional contributors of 1-hydroxypyrene among children near a steel mill. Cancer Epidemiol Biomarkers Prev. 18(1):96-101.

312. Lee MW, Chen ML, Lung SC, Tsai CJ, Lai CF, Yang SC, Mao IF. 2012. Increase of urinary concentrations of 8-hydroxy-2'-deoxyguanosine in diesel exhaust emission inspector exposed to polycyclic aromatic hydrocarbons. Int Arch Occup Environ Health. 85(3):273-82.

313. Lee SL, Wong WH, Lau YL. 2006. Association between air pollution and asthma admission among children in Hong Kong. Clin Exp Allergy 36:1138-1146.

314. Lee Y-L, Gilliland FD, Wang J-Y, Lee Y-C, Guo YL. 2008. Associations of FcepsilonRIbeta E237G polymorphism with wheezing in Taiwanese schoolchildren. Clinical and Experimental Allergy 38(3):413–420.

315. Lee YL, Hsiue TR, Lee YC, Lin YC, Guo YL. 2005. The association between glutathione S-transferase P1, M1 polymorphisms and asthma in Taiwanese schoolchildren. Chest. 128 (3):1156– 1162

316. Lee Y-L, Lin Y-C, Lee Y-C, Wang J-Y, Hsiue T-R, Guo YL. 2004. Glutathione S-transferase P1 gene polymorphism and air pollution as interactive risk factors for childhood asthma. Clinical and Experimental Allergy 34(11):1707–1713.

317. Lee Y-L, Shaw C-K, Su H-J, Lai J-S, Ko Y-C, Huang S-L, Sung F-C, Guo YL. 2003. Climate, traffic-related air pollutants and allergic rhinitis prevalence in middle-school children in Taiwan. European Respiratory Journal 21(6):964–970.

318. Li N, Alam J, Venkatesan MI, Schmitz D, Stefano E Di, Slaughter N, Killeen E, Wang X, Wang M, Miguel AH, Cho A, Sioutas C, Nel AE, Eiguren-fernandez A. 2004. Nrf2 is a key transcription factor that regulates antioxidant defense in macrophages and epithelial cells: protecting against the proinflammatory and oxidizing effects of diesel exhaust chemicals. The Journal of Immunology 173: 3467–3481.

319. Li N, Hao M, Phalen RF, Hinds WC, Nel AE. 2003a. Particulate air pollutants and asthma: A paradigm for the role of oxidative stress in PM-induced adverse health effects. Clinical Immunology 109(3):250–265.

320. Li N, Sioutas C, Cho A, Schmitz D, Misra C, Sempf J, Wang M, Oberley T, Froines J, Nel A. 2003b. Ultrafine Particulate Pollutants Induce Oxidative Stress and Mitochondrial Damage. Environmental Health Perspectives 111(4):455–460.

321. Li R, Shugart YY, Zhou W, An Y, Yang Y, Zhou Y, Zhang B, Lu D, Wang H, Qian J, Jin L. 2009. Common genetic variations of the cytochrome P450 1A1 gene and risk of hepatocellular carcinoma in a Chinese population. European Journal of Cancer 45: 1239–1247.

322. Li S, Williams G, Jalaludin B, Baker P. 2012. Panel studies of air pollution on children's lung function and respiratory symptoms: a literature review. J Asthma. 49(9):895-910.

323. Li Y-F, Gauderman WJ, Avol E, Dubeau L, Gilliland FD. 2006. Associations of tumor necrosis factor G-308A with childhood asthma and wheezing. American Journal of Respiratory and Critical Care Medicine 173(9):970–976.

324. Li YF, Tseng PJ, Lin CC, Hung CL, Lin SC, Su WC, Huang YL, Sung FC, Tai CK. 2009. NAD(P)H: Quinone oxidoreductase 1, glutathione S-transferase M1, environmental tobacco smoke exposure, and childhood asthma. Mutat Res. 678:53–58.

325. Lindgren A, Stroh E, Montnémery P, Nihlén U, Jakobsson K, Axmon A. 2009. Traffic-related air pollution associated with prevalence of asthma and COPD/chronic bronchitis. A cross-sectional study in Southern Sweden. International Journal of Health Geographics 8: 2.

326. Loft S, Poulsen HE. 1998. Estimation of oxidative DNA damage in man from urinary excretion of repair products. Acta Biochim Pol. 45(1):133-44.

327. London S, Romieu I. 2009. Gene by Environment Interaction in Asthma. Annual Review of Public Health. 30: 55-80.

328. Long J-R, Egan KM, Dunning L, Shu X-O, Cai Q, Cai H, Dai Q, Holtzman J, Gao Y-T, Zheng W. 2006. Population-based case-control study of AhR (aryl hydrocarbon receptor) and CYP1A2 polymorphisms and breast cancer risk. Pharmacogenetics and Genomics 16(4):237–243.

329. Lu SC. 1999. Regulation of hepatic glutathione synthesis: current concepts and controversies. FASEB J 13: 1169–1183.

330. Lurie N, Dubowitz T. 2007. Health disparities and access to health. JAMA 297(10):1118–1121.

331. Ma J, Betts NM. 2000. Zinc and copper intakes and their major food sources for older adults in the 1994-96 continuing survey of food intakes by individuals (CSFII). The Journal of Nutrition 130: 2838–2843.

332. Ma Q, Lu AYH. 2007. CYP1A induction and human risk assessment: an evolving tale of in vitro and in vivo studies. Drug Metabolism and Disposition 35: 1009–1016.

333. Maier A, Micka J, Miller K, Denko T, Chang CY, Nebert DW, Alvaro Puga. 1998. Aromatic hydrocarbon receptor polymorphism: development of new methods to correlate genotype with phenotype. Environmental Health Perspectives 106: 421–426.

334. Manini P, Palma G De, Andreoli R, Poli D, Mozzoni P, Folesani G, Mutti A, Apostoli P. 2006. Environmental and biological monitoring of benzene exposure in a cohort of Italian taxi drivers. Toxicology Letters 167(2):142–51.

335. Marczak ED, Marzec J, Zeldin DC, Kleeberger SR, Brown NJ, Pretorius M, Lee CR. 2012. Polymorphisms in the transcription factor NRF2 and forearm vasodilator responses in humans. Pharmacogenetics and Genomics 22(8):620–628.

336. Marzec JM, Christie JD, Reddy SP, Jedlicka AE, Vuong H, Lanken PN, Aplenc R, Yamamoto T, Yamamoto M, Cho H-Y, Kleeberger SR. 2007. Functional polymorphisms in the transcription factor NRF2 in humans increase the risk of acute lung injury. FASEB J 21(9):2237–2246.

337. Masuko H, Sakamoto T, Kaneko Y, Iijima H, Naito T, Noguchi E, Hirota T, Tamari M, Hizawa N.. 2011. Lower FEV1 in non-COPD, nonasthmatic subjects: association with smoking, annual decline in FEV1, total IgE levels, and TSLP genotypes. International Journal of Chronic Obstructive Pulmonary Disease. 6:181–189.

338. Masuko H, Sakamoto T, Kaneko Y, Iijima H, Naito T, Noguchi E, Hirota T, Tamari M, Hizawa N. 2011. An interaction between Nrf2 polymorphisms and smoking status affects annual decline in FEV_1: a longitudinal retrospective cohort study. BMC Medical Genetics 12: 97.

339. Martineau AR, Timms PM, Bothamley GH, Hanifa Y, Islam K, Claxton AP, Packe GE, Moore-Gillon JC, Darmalingam M, Davidson RN, Milburn HJ, Baker LV, Barker RD, Woodward NJ, Venton TR, Barnes KE, Mullett CJ, Coussens AK, Rutterford CM, Mein CA, Davies GR, Wilkinson RJ, Nikolayevskyy V, Drobniewski FA, Eldridge SM, Griffiths CJ. 2011. High-dose vitamin D(3) during intensive-phase antimicrobial treatment of pulmonary tuberculosis: a double-blind randomised controlled trial. Lancet. 377(9761):242-50.

340. Matés JM. 2000. Effects of antioxidant enzymes in the molecular control of reactive oxygen species toxicology. Toxicology 153: 83–104.

341. McConnell R, Berhane K, Yao L, Jerrett M, Lurmann F, Gilliland F, Künzli N, Gauderman J, Avol E, Thomas D, Peters J. 2006. Traffic, susceptibility, and childhood asthma. Environ Health Perspect. 114(5):766-72.

342. McConnell R, Islam T, Shankardass K, Jerrett M, Lurmann F, Gilliland F, Gauderman J, Avol E, Künzli N, Yao L, Peters J, Berhane K. 2011. Childhood incident asthma and traffic-related air pollution at home and school. Am J Respir Crit Care Med. 184(7):822-7.

343. McCreanor J, Cullinan P, Nieuwenhuijsen MJ, Stewart-Evans J, Malliarou E, Jarup L, Harrington R, Svartengren M, Han I-K, Ohman-Strickland P, Chung KF, Zhang J. 2007. Respiratory effects of exposure to diesel traffic in persons with asthma. The New England Journal of Medicine. 357(23):2348–2358.

344. McIlwain C., Townsend D., Tew K. 2006. Glutathione S-transferase polymorphisms: cancer incidence and therapy. Oncogene. 25:1639–1648.

345. Melén E, Nyberg F, Lindgren CM, Berglind N, Zucchelli M, Nordling E, Hallberg J, Svartengren M, Morgenstern R, Kere J, Bellander T, Wickman M, Pershagen G. 2008. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. Environmental Health Perspectives 116(8):1077–1084.

346. Mertz W. 1993. Chromium in human nutrition: a review. The Journal of Nutrition 123: 626–633.

347. Mesaros C, Arora JS, Wholer A, Vachani A, Blair IA. 2012. 8-Oxo-2'-deoxyguanosine as a biomarker of tobacco-smoking-induced oxidative stress. Free Radic Biol Med. 53(3):610-7.

348. Migliore E, Berti G, Galassi C, Pearce N, Forastiere F, Calabrese R, et al. 2009. Respiratory symptoms in children living near busy roads and their relationship to vehicular traffic: results of an Italian multicenter study (SIDRIA 2). Environmental Health 8: 27.

349. Minelli C, Granell R, Newson R, Rose-Zerilli MJ, Torrent M, Ring SM, Holloway JW, Shaheen SO, Henderson JA. 2010. Glutathione-S-transferase genes and asthma phenotypes: a Human Genome Epidemiology (HuGE) systematic review and meta-analysis including unpublished data. Int J Epidemiol. 39:539–562.

350. Minelli C, Wei I, Sagoo G, Jarvis D, Shaheen S, Burney P. 2011. Interactive effects of antioxidant genes and air pollution on respiratory function and airway disease: a HuGE review. Am J Epidemiol. 173(6):603-20.

351. Mimura T, Usui T, Mori M, Aixinjueluo W, Funatsu H, Noma H, Amano S. 2010. Immunochromatographic assay for measurement of total IgE in tears, nasal mucus, and saliva of patients with allergic rhinoconjunctivitis. J Asthma. 47(10):1153-60

352. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. 2010. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 363:1211-1221.

353. Møller P, Loft S. 2010. Oxidative damage to DNA and lipids as biomarkers of exposure to air pollution. Environ Health Perspect. 118(8):1126-36.

354. Morgenstern V, Zutavern A, Cyrys J, Brockow I, Gehring U, Koletzko S, Bauer CP, Reinhardt D, Wichmann H-E, Heinrich J. 2007. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occupational and Environmental Medicine 64(1):8–16.

355. Morgenstern V, Zutavern A, Cyrys J, Brockow I, Koletzko S, Krämer U, Behrendt H, Herbarth O, Berg A von, Bauer CP, Wichmann H-E, Heinrich J. 2008. Atopic diseases,

allergic sensitization, and exposure to traffic-related air pollution in children. American Journal of Respiratory and Critical Care Medicine 177(12):1331–1337.

356. Moshammer H, Hoek G, Luttmann-Gibson H, Neuberger MA, Antova T, Gehring U, Hruba F, Pattenden S, Rudnai P, Slachtova H, et al. 2006. Parental smoking and lung function in children: an international study. Am J Respir Crit Care Med. 173:1255–1263.

357. Moshammer H, Hutter HP, Hauck H, Neuberger M. 2006. Low levels of air pollution induce changes of lung function in a panel of schoolchildren. Eur Respir J. 27(6):1138-43.

358. Mudway IS, Behndig AF, Helleday R, Pourazar J, Frew AJ, Kelly FJ, Blomberg A. 2006. Vitamin supplementation does not protect against symptoms in ozone-responsive subjects. Free Radical Biology and Medicine 40(10):1702–1712.

359. Mulot C, Stücker I, Clavel J, Beaune P, Loriot M-A. 2005. Collection of Human Genomic DNA From Buccal Cells for Genetics Studies: Comparison Between Cytobrush, Mouthwash, and Treated Card. Journal of Biomedicine and Biotechnology 2005: 291–296.

360. Muñoz B, Magaña JJ, Romero-Toledo I, Juárez-Pérez E, López-Moya A, Leyva-García N, López-Campos C, Dávila-Borja VM, Albores A. 2012. The relationship among IL-13, GSTP1, and CYP1A1 polymorphisms and environmental tobacco smoke in a population of children with asthma in Northern Mexico. Environmental Toxicology and Pharmacology 33(2):226–232.

361. Naganuma T, Nakayama T, Sato N, Fu Z, Soma M, Aoi N, Hinohara S, Doba N, Usami R. 2008a. Association of extracellular superoxide dismutase gene with cerebral infarction in women: a haplotype-based case-control study. Hereditas 145(6):283–292.

362. Naganuma T, Nakayama T, Sato N, Fu Z, Soma M, Aoi N, Usami R. 2008b. A haplotype-based case-control study examining human extracellular superoxide dismutase gene and essential hypertension. Hypertension Research 31(8):1533–1540.

363. Nakai S, Nitta H, Maeda K. 1999. Respiratory health associated with exposure to automobile exhaust. Part III. Results of a cross-sectional study in 1987, and repeated pulmonary function tests from 1987 to 1990. Arch Environ Health 54:26–33.

364. Nakamura S, Kugiyama K, Sugiyama S, Miyamoto S, Koide S, Fukushima H, Honda O, Yoshimura M, Ogawa H. 2002. Polymorphism in the 5'-flanking region of human glutamate-cysteine ligase modifier subunit gene is associated with myocardial infarction. Circulation 105: 2968–2973.

365. Narayanan M, Owers-Bradley J, Beardsmore CS,MadaM, Ball I, Garipov R, Panesar KS, Kuehni CE, Spycher BD, Williams SE, et al. 2012. Alveolarization continues during

childhood and adolescence: new evidence from helium-3 magnetic resonance. Am J Respir Crit Care Med. 185:186–191.

366. Nebert DW, Roe AL, Vandale SE, Bingham E, Oakley GG. 2002. NAD(P)H:quinone oxidoreductase (NQO1) polymorphism, exposure to benzene, and predisposition to disease: a HuGE review. Genetics in Medicine 4(2):62–70.

367. Nebert DW, Ross A. 1996. Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. DNA Cell Biology 15: 273–280.

368. Nebert DW. 1989. The Ah locus: genetic differences in toxicity, cancer, mutation, and birth defects. Critical Reviews in Toxicology 20(3):153–174.

369. Neuberger M, Moshammer H, Kundi M. 2002. Declining ambient air pollution and lung function improvement in Austrian children. Atmos Environ 36:1733–6.

370. Ng CH, Janoo-Gilani R, Sipahimalani P, Gallagher RP, Gascoyne RD, Connors JM, Weber J-P, Lai AS, Leach S, Le ND, Brooks-Wilson AR, Spinelli JJ. 2010. Interaction between organochlorines and the AHR gene, and risk of non-Hodgkin lymphoma. Cancer Causes & Control 21(1):11–22.

371. Nicholson G, Smith A V., Jonsson F, Gustafsson O, Stefansson K, Donnelly P. 2002. Assessing population differentiation and isolation from single-nucleotide polymorphism data. Journal of the Royal Statistical Society - Series B: Statistical Methodology 64(4):695–715.

372. Nicolai T, Carr D, Weiland SK, Duhme H, von Ehrenstein O, Wagner C, von Mutius E. 2003. Urban traffic and pollutant exposure related to respiratory outcomes and atopy in a large sample of children. Eur Respir J 21:956–963.

373. Nicolai T. 2002. Pollution, environmental factors and childhood respiratory allergic disease. Toxicology 181-182: 317–21.

374. Nikasinovic L, Momas I, Seta N. 2003. Nasal epithelial and inflammatory response to ozone exposure: a review of laboratory-based studies published since 1985. Journal of Toxicology and Environmental Health. Part B, Critical reviews 6(5):521–568.

375. Nishita DM, Jack LM, McElroy M, McClure JB, Richards J, Swan GE, Bergen AW. 2009. Clinical trial participant characteristics and saliva and DNA metrics. BMC Medical Research Methodology 971.

376. Noakes PS, Thomas R, Lane C, Mori TA, Barden AE, Devadason SG, Prescott SL. 2007. Association of maternal smoking with increased infant oxidative stress at 3 months of age. Thorax. 62(8):714-7.

377. Nordenhäll C, Pourazar J, Ledin MC, Levin JO, Sandström T, Adelroth E. 2001. Diesel exhaust enhances airway responsiveness in asthmatic subjects. The European Respiratory Journal 17: 909–915.

378. Nordling E, Berglind N, Melén E, Emenius G, Hallberg J, Nyberg F, Pershagen G, Svartengren M, Wickman M, Bellander T. 2008. Traffic-related air pollution and childhood respiratory symptoms, function and allergies. Epidemiology 19(3):401–8.

379. Novotna B, Topinka J, Solansky I, Chvatalova I, Lnenickova Z, Sram RJ. 2007. Impact of air pollution and genotype variability on DNA damage in Prague policemen. Toxicology Letters 172(1-2):37–47.

380. Oftedal B, Brunekreef B, Nystad W, Madsen C, Walker SE, Nafstad P. 2008. Residential outdoor air pollution and lung function in schoolchildren. Epidemiology 19:129–137.

381. Oftedal B, Nystad W, Brunekreef B, Nafstad P. 2009. Long-Term Traffic-Related Exposures and Asthma Onset in Schoolchildren in Oslo, Norway. Environmental Health Perspectives 117: 839–844.

382. Okey AB, Vella LM, Harper PA. 1989. Detection and characterization of a low affinity form of cytosolic Ah receptor in livers of mice nonresponsive to induction of cytochrome P1-450 by 3-methylcholanthrene. Molecular Pharmacology 35: 823–830.

383. ONS. 2012. Population Estimates for England and Wales, Mid-2011 (2011 Census-based). Available: http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2011--2011-census-based-/index.html [accessed 13 April 2013].

384. Oosterlee A, Drijver M, Lebret E, Brunekreef B. 1996. Chronic respiratory symptoms in children and adults living along streets with high traffic density. Occup Environ Med 53:241-247.

385. Osburn WO, Kensler TW. 2008. Nrf2 signaling: an adaptive response pathway for protection against environmental toxic insults. Mutation Research 659: 31–39.

386. Otter M Von, Landgren S, Nilsson S, Celojevic D, Bergström P, Håkansson A, Nissbrandt H, Drozdzik M, Bialecka M, Kurzawski M, Blennow K, Nilsson M, Hammarsten O, Zetterberg H. 2010. Association of Nrf2-encoding NFE2L2 haplotypes with Parkinson's disease. BMC Medical Genetics 11(1):36.

387. Oury TD, Day BJ, Crapo JD. 1996. Extracellular superoxide dismutase in vessels and airways of humans and baboons. Free Radical Biology & Medicine 20: 957–965.

388. Pappas RS. 2011. Toxic elements in tobacco and in cigarette smoke: inflammation and sensitization. Metallomics. 3(11):1181-98.

389. Park SK, O'Neill MS, Wright RO, Hu H, Vokonas PS, Sparrow D, Suh H, Schwartz J. 2006. HFE genotype, particulate air pollution, and heart rate variability: a gene-environment interaction. Circulation. 114(25):2798-805.

390. Patel MM, Chillrud SN, Deepti KC, Ross JM, Kinney PL. 2013. Traffic-related air pollutants and exhaled markers of airway inflammation and oxidative stress in New York City adolescents. Environ Res. 121:71-8.

391. Pattenden S, Armstrong BG, Houthuijs D, Leonardi GS, Dusseldorp A, Boeva B, Hruba F, Brunekreef B, Fletcher T. Methodological approaches to the analysis of hierarchical studies of air pollution and respiratory health--examples from the CESAR study. 2000. Central European Study on Air pollution and Respiratory Health. J Expo Anal Environ Epidemiol. 10(5):420-6.

392. Pattenden S, Hoek G, Braun-Fahrlander C, Forastiere F, Kosheleva A, Neuberger M, et al. 2006. NO2 and children's respiratory symptoms in the PATY study. Occup Environ Med 63:828-835.

393. Pearson WR, Vorachek WR, Xu SJ, Berger R, Hart I, Vannais D, Patterson D. 1993. Identification of class-mu glutathione transferase genes GSTM1-GSTM5 on human chromosome 1p13. Am J Hum Genet. 53:220–233.

394. Pehkonen SO, Siefert R, Erel Y, Webb S, Hoffmann MR. 1993. Photoreduction of iron oxyhydroxides in the presence of important atmospheric organic compounds. Environmental Science & Technology 27(10):2056–2062.

395. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. 2005. Interpretative strategies for lung function tests. Eur Respir J 26:948-968.

396. Penard-Morand C, Raherison C, Charpin D, Kopferschmitt C, Lavaud F, Caillaud D, et al. 2010. Long-term exposure to close-proximity air pollution and asthma and allergies in urban children. Eur Respir J 36:33-40.

397. Peters JM, Avol E, Gauderman WJ, Linn WS, Navidi W, London SJ, Margolis H, Rappaport E, Vora H, Gong H Jr, Thomas DC. 1999. A study of twelve Southern California communities with differing levels and types of air pollution. II. Effects on pulmonary function. Am J Respir Crit Care Med. 159(3):768-75.

398. Piacentini S, Polimanti R, Moscatelli B, Re M a, Fuciarelli R, Manfellotto D, Fuciarelli M. 2010. Glutathione S-transferase gene polymorphisms and air pollution as interactive risk factors for asthma in a multicentre Italian field study: A preliminary study. Annals of Human Biology 37(3):427–439.

399. Pierse N, Rushton L, Harris RS, Kuehni CE, Silverman M, Grigg J. 2006. Locally generated particulate pollution and respiratory symptoms in young children. Thorax 61(3):216–220.

400. Pinkerton KE, Green FH, Saiki C, Vallyathan V, Plopper CG, Gopal V, Hung D, Bahne EB, Lin SS, Menache MG, et al. 2000. Distribution of particulate matter and tissue remodeling in the human lung. Environ Health Perspect 108:1063–1069.

401. Plopper CG, Weir AJ, Nishio SJ, Chang A, Voit M, Philpot RM, Buckpitt AR. 1994. Elevated susceptibility to 4-ipomeanol cytotoxicity in immature Clara cells of neonatal rabbits. The Journal of Pharmacology and Experimental Therapeutics 269: 867–880.

402. Plum LM, Rink L, Haase H. 2010. The Essential Toxin: Impact of Zinc on Human Health. International Journal of Environmental Research and Public Health 7: 1342–1365.

403. Poets CF, Stebbens VA, Samuels MP, Southall DP. 1993. Oxygen saturation and breathing patterns in children. Pediatrics 92: 686–690.

404. Pourazar J, Mudway IS, Samet JM, Helleday R, Wilson SJ, Frew AJ, Kelly FJ, Blomberg A, Sandstro T, Anthony J. 2009. Diesel exhaust activates redox-sensitive transcription factors and kinases in human airways. American Journal of Physiology Lung Cellular and Molecular Physiology 289(5):724–730.

405. Power MC, Weisskopf MG, Alexeeff SE, Wright RO, Coull BA, Spiro A 3rd, Schwartz J. 2013. Modification by hemochromatosis gene polymorphisms of the association between traffic-related air pollution and cognition in older men: a cohort study. Environ Health. 12:16.

406. Praud JP. 2004. Snoring in children: still many questions, only a few answers. Pediatr Pulmonol Suppl. 26:169-71.

407. Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–59.

408. Rahman I, Biswas SK, Kode A. 2006. Oxidant and antioxidant balance in the airways and airway diseases. European Journal of Pharmacology 533(1-3):222–239.#

409. Raizenne M, Neas LM, Damokosh AI, Dockery DW, Spengler JD, Koutrakis P, Ware JH, Speizer FE. 1996. Health effects of acid aerosols on North American children: pulmonary function. Environ Health Perspect. 104(5):506-14.

410. Rebbeck T. 1997. Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility. Cancer Epidemiol. Biomarkers Prev. 6:733–743.

411. Reddy NM, Kleeberger SR, Bream JH, Fallon PG, Kensler TW, Yamamoto M, Reddy SP. 2008. Genetic disruption of the Nrf2 compromises cell-cycle progression by impairing GSH-induced redox signaling. Oncogene 27: 5821–5832.

412. Reddy P, Naidoo ÃRN, Robins TG, Mentz G, Li H, London SJ. 2012. GSTM1 and GSTP1 Gene Variants and the Effect of Air Pollutants on Lung Function Measures in South African Children. American Journal of Industrial Medicine 55(12):1078–1086.

413. Reddy P, Naidoo RN, Robins TG, Mentz G, London SJ, Li H, Naidoo R. 2010. GSTM1, GSTP1, and NQO1 polymorphisms and susceptibility to atopy and airway hyperresponsiveness among South African schoolchildren. Lung. 188:409–414.

414. Reddy SP. 2008. The antioxidant response element and oxidative stress modifiers in airway diseases. Current Molecular Medicine 8: 376–383.

415. Reid L. 1977. 1976 Edward B.D. Neuhauser lecture: the lung: growth and remodeling in health and disease. American Journal of Roentgenology 129: 777–788.

416. Ren C, Fang S, Wright RO, Suh H, Schwartz J. 2011. Urinary 8-hydroxy-2'-deoxyguanosine as a biomarker of oxidative DNA damage induced by ambient pollution in the Normative Aging Study. Occupational and Environmental Medicine 68: 562–569.

417. Ren C, Vokonas PS, Suh H, Fang S, Christiani DC, Schwartz J. 2010. Effect modification of air pollution on Urinary 8-Hydroxy-2'-Deoxyguanosine by genotypes: an application of the multiple testing procedure to identify significant SNP interactions. Environmental Health 9(1):78.

418. Rengarajan K, Cristol SM, Mehta M, Nickerson JM. 2002. Quantifying DNA concentrations using fluorometry: a comparison of fluorophores. Molecular Vision 8: 416–421.

419. Renzetti G, Silvestre G, Amario CD, Bottini E, Bottini N, Auais A, Perez MK, Gloria-bottini F. 2009. Less Air Pollution Leads to Rapid Reduction of Airway Inflammation and Improved Airway Function in Asthmatic Children. Pediatrics 123(3):1051–1058.

420. Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, Ohman-Strickland P, Hu M, Philipp C, Diehl SR, Lu SE, Tong J, Gong J, Thomas D, Zhu T, Zhang JJ. 2012. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. JAMA. 307(19):2068-78.

421. Richter PA, Bishop EE, Wang J, Swahn MH. 2009. Tobacco smoke exposure and levels of urinary metals in the U.S. youth and adult population: the National Health and Nutrition

Examination Survey (NHANES) 1999-2004. Int J Environ Res Public Health. 6(7):1930-46.

422. Riihimäki V, Aitio A. 2012. Occupational exposure to aluminum and its biomonitoring in perspective. Crit Rev Toxicol. 42(10):827-53.

423. Robards K, Worsfold P. 1991. Cadmium: toxicology and analysis. A review. The Analyst 116: 549–68.

424. Rogan WJ. 1995. Environmental poisoning of children--lessons from the past. Environmental Health Perspectives 103: 19–23.

425. Rogers NL, Cole SA, Lan H-C, Crossa A, Demerath EW. 2007. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. American Journal of Human Biology 19: 319–326.

426. Rojas-Martinez R, Perez-Padilla R, Olaiz-Fernandez G, et al. 2007. Lung function growth in children with long-term exposure to air pollutants in Mexico city. Am J Respir Crit Care Med. 176:377–384.

427. Rolla G, Guida G, Heffler E, Badiu I, Bommarito L, De Stefani A, Usai A, Cosseddu D, Nebiolo F, Bucca C. 2007. Diagnostic classification of persistent rhinitis and its relationship to exhaled nitric oxide and asthma: a clinical study of a consecutive series of patients. Chest.131(5):1345-52.

428. Romieu I, Castro-Giner F, Kunzli N, Sunyer J. 2008. Air pollution, oxidative stress and dietary supplementation: a review. The European Respiratory Journal 31(1):179–197.

429. Romieu I, Ramirez-Aguilar M, Sienra-Monge JJ, Moreno-Macías H, Rio-Navarro BE del, David G, Marzec J, Hernández-Avila M, London S. 2006. GSTM1 and GSTP1 and respiratory health in asthmatic children exposed to ozone. The European Respiratory Journal 28(5):953–959.

430. Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M, Moreno-Macias H, Reyes-Ruiz NI, Estela del Rio-Navarro B, Hernandez-Avila M, London SJ. 2004. Genetic polymorphism of GSTM1 and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. Thorax. 59:8–10.

431. Rosenberg N a, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, Feldman MW. 2005. Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure. PLoS Genetics 1(6):e70.

432. Rosenlund M, Forastiere F, Porta D, Sario M De, Badaloni C, Perucci C. 2009. Traffic-related air pollution in relation to respiratory symptoms, allergic sensitisation and lung function in schoolchildren. Thorax 64(7):573–80.

433. Rosman, KJR, Taylor, PDP. 1999. Abundance Measurements. Pure Appl. Chem. 71, 1593-1607

434. Ross D, Kepa JK, Winski SL, Beall HD, Anwar A, Siegel D. 2000. NAD(P)H:quinone oxidoreductase 1 (NQO1): chemoprotection, bioactivation, gene regulation and genetic polymorphisms. Chemico Biological Interactions 129: 77–97.

435. Rossner P Jr, Rossnerova A, Spatova M, Beskid O, Uhlirova K, Libalova H, Solansky I, Topinka J, Sram RJ. 2013. Analysis of biomarkers in a Czech population exposed to heavy air pollution. Part II: chromosomal aberrations and oxidative stress. Mutagenesis. 28(1):97-106.

436. Rossner P Jr, Svecova V, Milcova A, Lnenickova Z, Solansky I, Santella RM, Sram RJ. 2007. Oxidative and nitrosative stress markers in bus drivers. Mutat Res. 617(1-2):23-32.

437. Rotunno M, Yu K, Lubin JH, Consonni D, Pesatori AC, Goldstein AM, Goldin LR, Wacholder S, Welch R, Burdette L, Chanock SJ, Bertazzi PA, Tucker MA, Caporaso NE, Chatterjee N, Bergen AW, Landi MT. 2009. Phase I Metabolic Genes and Risk of Lung Cancer: Multiple Polymorphisms and mRNA Expression. PLoS ONE 4: 13.

438. Roy A, Chapman RS, Hu W, Wei F, Liu X, Zhang J. 2012a. Indoor air pollution and lung function growth among children in four Chinese cities. Indoor Air. 22(1):3-11.

439. Roy A, Hu W, Wei F, Korn L, Chapman RS, Zhang JJ. 2012b. Ambient particulate matter and lung function growth in Chinese children. Epidemiology. 23(3):464-72.

440. Ryan PH, LeMasters G, Biagini J, Bernstein D, Grinshpun S, Shukla R, Wilson K, Villareal M, Burkle J, Lockey J. 2005. Is it traffic type, volume, or distance? Wheezing in infants living near truck and bus traffic. The Journal of Allergy and Clinical Immunology 116(2):279–284.

441. Ryan PH, Lemasters GK, Biswas P, Levin L, Hu S, Lindsey M, Bernstein DI, Lockey J, Villareal M, Khurana Hershey GK, Grinshpun S. 2007. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. Environmental Health Perspectives 115(2):278–284.

442. Rylander-Rudqvist T, Håkansson N, Tybring G, Wolk A. 2006. Quality and quantity of saliva DNA obtained from the self-administrated oragene method--a pilot study on the cohort of Swedish men. Cancer Epidemiology, Biomarkers & Prevention 15(9):1742–1745.

443. Masuko H, Sakamoto T, Kaneko Y, Iijima H, Naito T, Noguchi E, Hirota T, Tamari M, Hizawa N.2011. Lower FEV1 in non-COPD, nonasthmatic subjects: association with smoking, annual decline in FEV1, total IgE levels, and TSLP genotypes. International Journal of Chronic Obstructive Pulmonary Disease. 6:181–189.

444. Salam MT, Lin P-C, Avol EL, Gauderman WJ, Gilliland FD. 2007. Microsomal epoxide hydrolase, glutathione S-transferase P1, traffic and childhood asthma. Thorax 62(12):1050–1057.

445. Salam MT, Byun HM, Lurmann F, Breton CV, Wang X, Eckel SP, Gilliland FD. 2012. Genetic and epigenetic variations in inducible nitric oxide synthase promoter, particulate pollution, and exhaled nitric oxide levels in children. J Allergy Clin Immunol. 129(1):232-9.e1-7.

446. Salam MT, Bastain TM, Rappaport EB, Islam T, Berhane K, Gauderman WJ, Gilliland FD. 2011. Genetic variations in nitric oxide synthase and arginase influence exhaled nitric oxide levels in children. Allergy. 66(3):412-9.

447. Salvi S, Blomberg A, Rudell B, Kelly F, Sandström T, Holgate ST, Frew A. 1999. Acute inflammatory responses in the airways and peripheral blood after short-term exposure to diesel exhaust in healthy human volunteers. American Journal of Respiratory and Critical Care Medicine 159: 702–709.

448. Samoli E, Nastos PT, Paliatsos AG, Katsouyanni K, Priftis KN. 2011. Acute effects of air pollution on pediatric asthma exacerbation: evidence of association and effect modification. Environ Research 111:418-424.

449. San Jose C, Cabanillas A, Benitez J, Carrillo JA, Jimenez M, Gervasini G. 2010. CYP1A1 gene polymorphisms increase lung cancer risk in a high-incidence region of Spain: a case control study. BMC Cancer 10: 463.

450. Sandström J, Nilsson P, Karlsson K, Marklund SL. 1994. 10-fold increase in human plasma extracellular superoxide dismutase content caused by a mutation in heparin-binding domain. The Journal of Biological Chemistry 269: 19163–19166.

451. Saoudi A, Zeghnoun A, Bidondo ML, Garnier R, Cirimele V, Persoons R, Fréry N. 2012. Urinary arsenic levels in the French adult population: the French National Nutrition and Health Study, 2006-2007. Sci Total Environ. 433:206-15.

452. Sasaki H, Suzuki A, Shitara M, Hikosaka Y, Okuda K, Moriyama S, Yano M, Fujii Y. 2013. Polymorphisms of NRF2 gene correlated with decreased FEV1 in lung cancers of smokers. Biomedical Reports 1484–488.

453. Schroeder HA, Balassa JJ, Tipton IH. 1970. Essential trace metals in man: Molybdenum. Journal of Chronic Diseases 23: 481–499.

454. Schroer KT, Biagini Myers JM, Ryan PH, LeMasters GK, Bernstein DI, Villareal M, Lockey JE, Reponen T, Grinshpun S, Khurana Hershey GK. 2009. Associations between

multiple environmental exposures and Glutathione S-Transferase P1 on persistent wheezing in a birth cohort. The Journal of Pediatrics 154(3):401–408.

455. Schultz E, Gruzieva O, Bellander T, Bottai M, Hallberg G, Kull I, Melen E. 2012. Traffic-related air pollution and lung function in children at 8 years of age: a birth cohort study. Am J Respir Crit Care Med. 186: 1286–1291.

456. Schwartz J. Air pollution and children's health. Pediatrics. 2004; 113(suppl):1037–1043.

457. Seet RC, Lee CY, Loke WM, Huang SH, Huang H, Looi WF, Chew ES, Quek AM, Lim EC, Halliwell B. 2011. Biomarkers of oxidative damage in cigarette smokers: which biomarkers might reflect acute versus chronic oxidative stress? Free Radic Biol Med. 50(12):1787-93.

458. Sehlstedt M, Dove R, Boman C, Pagels J, Swietlicki E, Löndahl J, Westerholm R, Bosson J, Barath S, Behndig AF, Pourazar J, Sandström T, Mudway IS, Blomberg A. 2010. Antioxidant airway responses following experimental exposure to wood smoke in man. Particle and Fibre Toxicology 721.

459. Sekine K, Shima M, Nitta Y, Adachi M. 2004. Long term effects of exposure to automobile exhaust on the pulmonary function of female adults in Tokyo, Japan. Occup Environ Med 61:350–357.

460. Shen R, Fan J-B, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, Steemers F, Garcia F, Kermani BG, Gunderson K, Oliphant A. 2005. High-throughput SNP genotyping on universal bead arrays. Mutation Research 573(1-2):70–82.

461. Shi MM, Kugelman A, Iwamoto T, Tian L, Forman HJ. 1994. Quinone-induced oxidative stress elevates glutathione and induces gamma-glutamylcysteine synthetase activity in rat lung epithelial L2 cells. The Journal of Biological Chemistry 269: 26512–26517.

462. Shinyashiki M, Eiguren-Fernandez A, Schmitz DA, Stefano E Di, Li N, Linak WP, Cho S-H, Froines JR, Cho AK. 2009. Electrophilic and redox properties of diesel exhaust particles. Environmental Research 109: 239–244.

463. Shimbo S, Watanabe T, Nakatsuka H, Yaginuma-Sakurai K, Ikeda M. 2013. Dietary tin intake and association with canned food consumption in Japanese preschool children. Environ Health Prev Med. 18(3):230-6.

464. Siedlinski M, Diemen CC van, Postma DS, Vonk JM, Boezen HM. 2009a. Superoxide dismutases, lung function and bronchial responsiveness in a general population. The European Respiratory Journal 33(5):986–992.

465. Siedlinski M, Postma DS, Boer JM, Steege G Van Der, Schouten JP, Smit HA, Boezen HM. 2009b. Level and course of FEV1 in relation to polymorphisms in NFE2L2 and KEAP1 in the general population. Respiratory Research 10: 73.

466. Sies H. 1991. Oxidative Stress: From Basic Research to Clinical Application. The American Journal of Medicine 91: 31–38.

467. Sigma Technical Notes. 2007. Whole Genome Amplification Advisor.

468. Simes RJ. 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika 73(3):751–754.

469. Singh A, Rangasamy T, Thimmulappa RK, Lee H, Osburn WO, Brigelius-Flohé R, Kensler TW, Yamamoto M, Biswal S. 2006. Glutathione Peroxidase 2, the Major Cigarette Smoke–Inducible Isoform of GPX in Lungs, Is Regulated by Nrf2. American Journal of Respiratory Cell and Molecular Biology 35: 639–650.

470. Singh B, Chatterjee A, Ronghe AM, Bhat NK, Bhat HK. 2013. Antioxidant-mediated up-regulation of OGG1 via NRF2 induction is associated with inhibition of oxidative DNA damage in estrogen-induced breast cancer. BMC Cancer. 13:253.

471. Smargiassi A, Berrada K, Fortier I, Kosatsky T. 2006. Traffic intensity, dwelling value, and hospital admissions for respiratory disease among the elderly in Montreal (Canada): a case-control analysis. J Epidemiol Community Health 60(6):507–512.

472. Smart J, Daly a K. 2000. Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. Pharmacogenetics 10: 11–24.

473. Sonnappa S, Bastardo CM, Aurora P, Whyatt D, Kelly FJ, Beevers SD, Grigg J. 2011. Residential proximity to main roads and lung function in healthy young children. Eur Respir J. 37(6):1525-6.

474. Sørensen M, Autrup H, Møller P, Hertel O, Jensen SS, Vinzents P, Knudsen LE, Loft S. 2003. Linking exposure to environmental pollutants with biological effects. Mutat Res. 544(2-3):255-71.

475. Sørheim IC, DeMeo DL, Washko G, Litonjua A, Sparrow D, Bowler R, Bakke P, Pillai SG, Coxson HO, Lomas DA, Silverman EK, Hersh CP. 2010. Polymorphisms in the superoxide dismutase-3 gene are associated with emphysema in COPD. COPD 7: 262–268.

476. Stanojevic S, Wade A, Cole TJ, Lum S, Custovic A, Silverman M, Hall GL, Welsh L, Kirkby J, Nystad W, Badier M, Davis S, Turner S, Piccioni P, Vilozni D, Eigen H, Vlachos-Mayer H, Zheng J, Tomalak W, Jones M, Hankinson JL, Stocks J; Asthma UK Spirometry Collaborative Group. 2009. Spirometry centile charts for young Caucasian

children: the Asthma UK Collaborative Initiative. Am J Respir Crit Care Med. 180(6):547-52.

477. Stern DA, Morgan WJ, Wright AL, Guerra S, Martinez FD. 2007. Poor airway function in early infancy and lung function by age 22 years: a non-selective longitudinal cohort study. Lancet 370:758–64.

478. Stohs SJ, Bagchi D, Bagchi M. 1997. Toxicity of trace elements in tobacco smoke. Inhalation Toxicology 9: 867–890.

479. Strachan DP, Anderson HR, Limb ES, O'Neill A, Wells N. 1994. A national survey of asthma prevalence, severity, and treatment in Great Britain. Arch Dis Child 70:174-178.

480. Strange RC, Spiteri MA, Ramachandran S, Fryer AA. 2001. Glutathione-S-transferase family of enzymes. Mutat Res. 482:21–26.

481. Su WY, Folz R, Chen JS, Crapo JD, Chang LY. 1997. Extracellular superoxide dismutase mRNA expressions in the human lung by in situ hybridization. American Journal of Respiratory Cell and Molecular Biology 16: 162–170.

482. Svecova V, Rossner P Jr, Dostal M, Topinka J, Solansky I, Sram RJ. 2009. Urinary 8-oxodeoxyguanosine levels in children exposed to air pollutants. Mutat Res. 662(1-2):37-43.

483. Taioli E, Ford J, Trachman J, Li Y, Demopoulos R, Garte S. 1998. Lung cancer risk and CYP1A1 genotype in African Americans. Carcinogenesis 19: 813–817.

484. Tamer L, Calikoglu M, Ates NA, Yildirim H, Ercan B, Saritas E, Unlu A, Atik U. 2004. Glutathione-S-transferase gene polymorphisms (GSTT1, GSTM1, GSTP1) as increased risk factors for asthma. Respirology. 9:493–498.

485. Teo YY, Inouye M, Small KS, Gwilliam R, Kwiatkowski DP, Clark TG. 2009. A genotype calling algorithm for the Illumina BeadArray platform. Bioinformatics 23(20):2741–2746.

486. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. 1998. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. Lancet 351:1225-1232.

487. Thomas NS, Wilkinson J, Holgate ST. 1997. The candidate region approach to the genetics of asthma and allergy. American Journal of Respiratory and Critical Care Medicine 156: S144–S151.

488. Thorisson GA, Smith AV, Krishnan L, Stein LD. 2005. The international HapMap project web site. Genome Res 15:1592–1593.

489. Thorpe A, Harrison RM. 2008. Sources and properties of non-exhaust particulate matter from road traffic: a review. Sci Total Environ. 400(1-3):270-82.

490. Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, Klareskog L, Pulver AE, Qi L, Gregersen PK, et al. 2008. Analysis and application of European genetic substructure using 300 K SNP information. PLoS Genet. 2008;4(1):e4.

491. Tintle N, Gordon D, Bruggen D Van, Finch S. 2009. The cost effectiveness of duplicate genotyping for testing genetic association. Annals of Human Genetics 73: 370–378.

492. Topinka J, Sevastyanova O, Binkova B, Chvatalova I, Milcova A, Lnenickova Z, Novakova Z, Solansky I, Sram RJ. 2007. Biomarkers of air pollution exposure--a study of policemen in Prague. Mutation Research 624(1-2):9–17.

493. Tosic M, Ott J, Barral S, Bovet P, Deppen P, Gheorghita F, Matthey ML, Parnas J, Preisig M, Saraga M, Solida A, Timm S, Wang AG, Werge T, Cuénod M, Do KQ. 2006. Schizophrenia and Oxidative Stress: Glutamate Cysteine Ligase Modifier as a Susceptibility Gene. Am J Hum Genet. 2006 Sep;79(3):586-92.

494. Totlandsdal AI, Cassee FR, Schwarze P, Refsnes M, Låg M. 2010. Diesel exhaust particles induce CYP1A1 and pro-inflammatory responses via differential pathways in human bronchial epithelial cells. Particle and Fibre Toxicology 7(1):41.

495. Tower Hamlets Council. 2011. Asthma: factsheet, Tower Hamlets Joint Strategic Needs Assessment 2010-2011. Available: http://www.towerhamlets.gov.uk/search.aspx?cx=008280465879053608327:251xeiiz6ey&cof=FORID:11;NB:1&ie=UTF-8&q=asthma%20factsheet.

496. Townsend DM, Tew KD, Tapiero H. 2003. The importance of glutathione in human disease. Biomedicine & Pharmacotherapy 57145–155.

497. Transport for London. 2008. London Low Emission Zone Impacts Monitoring Baseline Report, July 2008. London:Transport for London.

498. Tujague J, Bastaki M, Holland N, Balmes JR, Tager IB. 2006. Antioxidant intake, GSTM1 polymorphism and pulmonary function in healthy young adults. Eur Respir J. 27:282–288.

499. Tung K-Y, Tsai C-H, Lee YL. 2011. Microsomal epoxide hydroxylase genotypes/diplotypes, traffic air pollution, and childhood asthma. Chest 139(4):839–48.

500. Turnovska TH, Marinov BI. 2009. The influence of air pollution during intrauterine development and early childhood on respiratory functions at later age. Int J Hyg Environ Health 212:519–532.

501. Ungvári I, Hadadi E, Virág V, Nagy A, Kiss A, Kalmár A, Zsigmond G, Semsei AF, Falus A, Szalai C. 2012. Relationship between air pollution, NFE2L2 gene polymorphisms and childhood asthma in a Hungarian population. Journal of Community Genetics 3(1):25–33.

502. Valavanidis A, Fiotakis K, Vlachogianni T. 2008. Airborne particulate matter and human health: toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms. Journal of Environmental Science and Health Part C 26(4):339–362.

503. Valavanidis A, Vlachogianni T, Fiotakis C. 2009. 8-hydroxy-2' -deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis. Journal of Environmental Science and Health Part C 27(2):120–139.

504. van der Zee S, Hoek G, Boezen HM, Schouten JP, van Wijnen JH, Brunekreef B. 1999. Acute effects of urban air pollution on respiratory health of children with and without chronic respiratory symptoms. Occup Environ Med 56:802-812.

505. van Erp AM, O'Keefe R, Cohen AJ, Warren J. 2008. Evaluating the effectiveness of air quality interventions. J Toxicol Environ Health A 71:583–587.

506. von Otter M, Landgren S, Nilsson S, Celojevic D, Bergström P, Håkansson A, Nissbrandt H, Drozdzik M, Bialecka M, Kurzawski M, Blennow K, Nilsson M, Hammarsten O, Zetterberg H. 2010. Association of NRF2-encoding NFE2L2 haplotypes with Parkinson's disease. BMC Medical Genetics. 2;11:36.

507. Vawda S, Mansour R, Takeda A, Funnell P, Kerry S, Mudway I, Jamaludin J, Shaheen S, Griffiths C, Walton R. 2013. Associations between inflammatory and immune response genes and adverse respiratory outcomes following exposure to outdoor air pollution: a HuGE systematic review. Am J Epidemiol. 179(4):432-42.

508. Veglia F, Loft S, Matullo G, Peluso M, Munnia A, Perera F, Phillips DH, Tang D, Autrup H, Raaschou-Nielsen O, Tjønneland A, Vineis P. 2008. DNA adducts and cancer risk in prospective studies: a pooled analysis and a meta-analysis. Carcinogenesis 29(5):932–936.

509. Venn A, Lewis S, Cooper M, Hubbard R, Hill I, Boddy R, Bell M, Britton J. 2000. Local road traffic activity and the prevalence, severity, and persistence of wheeze in school children: combined cross sectional and longitudinal study. Occupational and Environmental Medicine 57: 152–158.

510. Vogel CF, Sciullo E, Wong P, Kuzmicky P, Kado N, Matsumura F. 2005. Induction of proinflammatory cytokines and C-reactive protein in human macrophage cell line U937 exposed to air pollution particulates.Environ Health Perspect. 113:1536-1541.

511. Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. 2010. Interpretation of Association Signals and Identification of Causal Variants from Genome-wide Association Studies. The American Journal of Human Genetics 86: 730–742.

512. Wang S, Chanock S, Tang D, Li Z, Jedrychowski W, Perera FP. 2008. Assessment of interactions between PAH exposure and genetic polymorphisms on PAH-DNA adducts in African American, Dominican, and Caucasian mothers and newborns. Cancer Epidemiology, Biomarkers & Prevention 17(2):405–413.

513. Watson M a, Stewart RK, Smith GB, Massey TE, Bell D. 1998. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. Carcinogenesis 19: 275–80.

514. Watters JL, Satia JA, Costa K-A Da, Boysen G, Collins LB, Morrow JD, Milne GL, Swenberg JA. 2009. Comparison of three oxidative stress biomarkers in a sample of healthy adults. Biomarkers 14: 587–595.

515. Watters JL, Satia JA, Kupper LL. 2008. Correlates of antioxidant nutrients and oxidative DNA damage differ by race in a cross-sectional study of healthy African American and white adults. Nutrition Research 28: 565–576.

516. Weiland SK, Mundt KA, Ruckmann A, Keil U. 1994. Self-reported wheezing and allergic rhinitis in children and traffic density on street of residence. Ann Epidemiol 4:243-247.

517. Weimann A, Broedbaek K, Henriksen T, Stovgaard ES, Poulsen HE. 2012. Assays for urinary biomarkers of oxidatively damaged nucleic acids. Free Radic Res. 46(4):531-40.

518. Weir A, Westerhoff P, Fabricius L, Hristovski K, Goetz N Von. 2012. Titanium dioxide nanoparticles in food and personal care products. Environmental Science & Technology 46(4):2242–50.

519. Wenten M, Li YF, Lin PC, Gauderman WJ, Berhane K, Avol E, Gilliland FD. 2009. In utero smoke exposure, glutathione S-transferase P1 haplotypes, and respiratory illness-related absence among schoolchildren. Pediatrics. 123(5):1344-51.

520. Whitelaw ML, Gustafsson JA, Poellinger L. 1994. Identification of transactivation and repression functions of the dioxin receptor and its basic helix-loop-helix/PAS partner factor Arnt: inducible versus constitutive modes of regulation. Molecular and Cellular Biology 14: 8343–8355.

521. Winkler CA, Nelson GW, Smith MW. 2010. Admixture mapping comes of age. Annu Rev Genomics Hum Genet. 11:65-89.

522. WHO. 1998. IPCS Environmental Health Criteria No. 200: Copper, World Health Organization, Geneva.

523. WHO. 2004. Health aspects of air pollution: Results from the WHO project “Systematic review of health aspects of air pollution in Europe”. Publications World Health Organization Regional Office for Europe.

524. WHO. 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide; Global update 2005; Summary of risk assessment. Available: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf.

525. Wigginton JE, Cutler DJ, Abecasis GR. 2005. A note on exact tests of Hardy-Weinberg equilibrium. The American Journal of Human Genetics 76(5):887–893.

526. Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW. 1997. Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. Proceedings of the National Academy of Sciences of the United States of America 94: 3195–3199.

527. Wilson DM 3rd, Kim D, Berquist BR, Sigurdson AJ. 2011. Variation in base excision repair capacity. Mutat Res. 711(1-2):100-12.

528. Wittke-Thompson JK, Pluzhnikov A, Cox NJ. 2005. Rational inferences about departures from Hardy-Weinberg equilibrium. The American Journal of Human Genetics 76(6):967–86.

529. Wjst M, Reitmeir P, Dold S, Wulff A, Nicolai T, von Loeffelholz-Colberg EF, von Mutius E. 1993. Road traffic and adverse effects on respiratory health in children. BMJ 307:596–600

530. Wong JM, Harper P a, Meyer U a, Bock KW, Morike K, Lagueux J, Ayotte P, Tyndale RF, Sellers EM, Manchester DK, Okey a B. 2001a. Ethnic variability in the allelic distribution of human aryl hydrocarbon receptor codon 554 and assessment of variant receptor function in vitro. Pharmacogenetics 11: 85–94.

531. Wong JM, Okey a B, Harper P a. 2001b. Human aryl hydrocarbon receptor polymorphisms that result in loss of CYP1A1 induction. Biochemical and Biophysical Research Communications 288(4):990–996.

532. Woodruff TJ, Parker JD, Schoendorf KC. 2006. Fine Particulate Matter (PM2.5) Air Pollution and Selected Causes of Postneonatal Infant Mortality in California. Environmental Health Perspectives 114(5):786–790.

533. Wright DT, Cohn LA, Li H, Fischer B, Li CM, Adler KB. 1994. Interactions of oxygen radicals with airway epithelium. Environmental Health Perspectives 102: 85–90.

534. Wright S. 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. Evolution 19: 395–420.

535. Wu P, Roberts LJ 2nd, Shintani AK, Sheller JR, Minton PA, Higgins SB, Hartert TV. 2007. Changes in urinary dinor dihydro F(2)-isoprostane metabolite concentrations, a

marker of oxidative stress, during and following asthma exacerbations. Free Radic Res. 41(9):956-62.

536. Wu W, Peden DB, McConnell R, Fruin S, Diaz-Sanchez D. 2012. Glutathione-S-transferase M1 regulation of diesel exhaust particle-induced pro-inflammatory mediator expression in normal human bronchial epithelial cells. Part Fibre Toxicol. 9:31

537. Wyler C, Braun-Fahrländer C, Künzli N, Schindler C, Ackermann-Liebrich U, Perruchoud AP, Leuenberger P, Wüthrich B. 2000. Exposure to motor vehicle traffic and allergic sensitization. The Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. Epidemiology 11: 450–456.

538. Xiao GG, Wang M, Li N, Loo J a, Nel AE. 2003. Use of proteomics to demonstrate a hierarchical oxidative stress response to diesel exhaust particle chemicals in a macrophage cell line. The Journal of Biological Chemistry 278(50):50781–50790.

539. Yang I a, Holz O, Jörres R a, Magnussen H, Barton SJ, Rodríguez S, Cakebread J a, Holloway JW, Holgate ST. 2005. Association of tumor necrosis factor-alpha polymorphisms and ozone-induced change in lung function. American Journal of Respiratory and Critical Care Medicine 171(2):171–176.

540. Yang Y, Dieter MZ, Chen Y, Shertzer HG, Nebert DW, Dalton TP. 2002. Initial characterization of the glutamate-cysteine ligase modifier subunit Gclm(-/-) knockout mouse. Novel model system for a severely compromised oxidative stress response. The Journal of Biological Chemistry 277: 49446–49452.

541. Young RP, Hopkins R, Black PN, Eddy C, Wu L, Gamble GD, Mills GD, Garrett JE, Eaton TE, Rees MI. 2006. Functional variants of antioxidant genes in smokers with COPD and in those with normal lung function. Thorax 61(5):394–399.

542. Yu J, Kang MJ, Kim BJ, Kwon JW, Song YH, Choi WA, et al. 2011. Polymorphisms in GSDMA and GSDMB are associated with asthma susceptibility, atopy and BHR. Pediatr Pulmonol 46:701-708.

543. Yu TS, Wong TW, Wang XR, Song H, Wong SL, Tang JL. 2001. Adverse effects of low-level air pollution on the respiratory health of schoolchildren in Hong Kong. J Occup Environ Med. 43:310–16.

544. Zanobetti A, Schwartz J. 2009. The Effect of Fine and Coarse Particulate Air Pollution on Mortality: A National Analysis. Environmental Health Perspectives 898(6):898–903.

545. Zhan P, Wang Q, Qian Q, Wei S-Z, Yu L-K. 2011. CYP1A1 MspI and exon7 gene polymorphisms and lung cancer risk: an updated meta-analysis and review. J Exp Clin Cancer Res. 30:99.

546. Zhang Y, Tao S, Shen H, Ma J. 2009. Inhalation exposure to ambient polycyclic aromatic hydrocarbons and lung cancer risk of Chinese population. Proceedings of the National Academy of Sciences of the United States of America 106: 21063–21067.

547. Zhang ZY, Fasco MJ, Huang L, Guengerich FP, Kaminsky LS. 1996. Characterization of purified human recombinant cytochrome P4501A1-Ile462 and -Val462: assessment of a role for the rare allele in carcinogenesis. Cancer Research 56: 3926–3933.

548. Zhu T, Melamed M, Parrish D, Gauss M, Klenner LG, Lawrence M, Konare A, Liousse C. 2012. WMO/IGAC impacts of megacities on air pollution and climate: World Meteorological Organization/Global Atmosphere Watch. Geneva, Switzerland: World Meteorological Organization; GAW Report No. 25.

549. Zhu Y, Hinds WC, Kim S, Sioutas C. 2002. Concentration and size distribution of ultrafine particles near a major highway. J Air Waste Manag Assoc 52:1032–1042.

550. Zubero MB, Aurrekoetxea JJ, Ibarluzea JM, Arenaza MJ, Rodríguez C, Sáenz JR. 2010. Heavy metal levels (Pb, Cd, Cr and Hg) in the adult general population near an urban solid waste incinerator. Sci Total Environ. 408(20):4468-74.

551. Zuraimi MS, Tham KW, Chew FT, Ooi PL, Koh D. 2011. Home air-conditioning, traffic exposure, and asthma and allergic symptoms among preschool children. Pediatr Allergy Immunol 22:e112-118.

Appendix A

Details of School Locations and Subject Recruitment in the London LEZ Schools Study

Figure A1 illustrates the location of the schools in Tower Hamlets and Hackney. The number of children recruited at each school is summarised in **Table A1** and the number of endpoint examined during the health assessments are reported in **Table A2**.

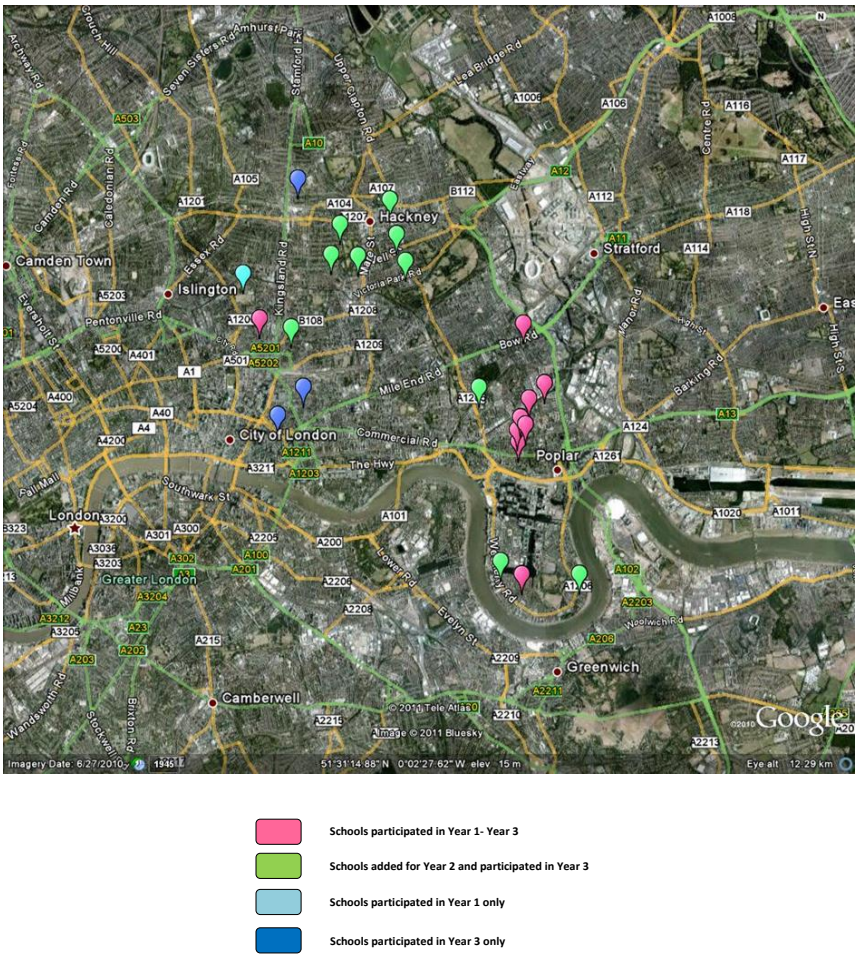


Figure A1 Map of the schools locations participating in the London LEZ Schools Study (Year 1-3).

Table A1 Participant numbers by school (Years 1-3)

School	Years 1-3 (2008-11)	Year 1 (2008-09)	Year 2 (2009-10)	Year 3 (2010-11)
Bygrove	50 (55)	6 (20)	29 (94)	15 (50)
Christ Church#	18 (75)			18 (75)
Colvestone#	14 (56)			14 (56)
Gayhurst*	80 (54)		39 (52)	41 (57)
Harbinger	47 (57)	20 (74)	14 (54)	13 (43)
Holy Family	53 (62)	22 (85)	14 (47)^	17 (57)
Lansbury Lawrence	109 (63)	37 (69)	41 (68)	31 (52)
Lauriston*	41 (69)		14 (48)	27 (90)
London Fields*	52 (61)		30 (57)	22 (69)
Manorfield	58 (39)	23 (40)	28 (47)	7 (23)
Mayflower	38 (42)	12 (40)	15 (50)	11 (37)
Morningside*	32 (38)		16 (30)	16 (57)
Orchard*	53 (49)		33 (61)	20 (36)
Sir John Cass#	27 (93)			27 (93)
St Agnes	65 (74)	23 (79)	21 (70)	21 (72)
St Edmund's*	20 (42)		13 (54)	7 (29)
St Luke's*	40 (68)		20 (67)	20 (69)
St Monica's	49 (86)	17 (59)	28 (93)	21 (78)
St Paul with St Luke*	21 (36)		14 (50)	7 (23)^
St Paul's with St Michael's*	23 (50)		11 (38)	12 (71)
St Saviour's	55 (63)	19 (63)	19 (66)	17 (61)
Virginia*	37 (73)		17 (77)	20 (69)
Whitmore	20 (40)	20 (40)		
TOTAL (%)	**1019 (56)**	**199 (55)**	**416 (57)**	**404 (56)**

Notes: figure in brackets indicate % of participants; * new school for Year 2; # new school for Year 3; ^ class size estimated at 30 as figure not provided by school.

Table A2 Details of data collected during health assessments each year

Details of Measurement	Year 1 (2008-09)		Year 2 (2009-10)		Year 3 (2010-11)		Year 1- 3 (2008-11)
	n	%	n	%	n	%	n
Participating schools	10		19		22		ǂ23
Participating children	199	55	416	57	404	56	1019
Height & weight measured	191	96	416	100	403	99.8	1010
Exhaled nitric oxide (eNO)	185	93	401	96	400	99	986
Post bronchodilator FEV1	150	75	402	97	399	99	951
Post bronchodilator FVC	148	74	384	92	388	96	920
Urine sample	178	89	398	96	398	98	973
Successfully DNA extracted (from saliva sample)	173	87	413	99	401	99	987
DNA samples passed genotyping QA/QC	170	85	406	98	398	99	974
Modelled air quality variables matched to home address	196	98	415	99.8	404	100	1015
Returned respiratory health questionnaire	129	65	383	92	389	96	901
Additional questionnaires (returned by children who declined to participate in health assessments)	3		36		55		94

ǂ Total number of different schools visited throughout Year 1- 3 (2008-11) study.

Appendix B

Preliminary Study on DNA Extraction and Quantification Optimization

For the preliminary optimization studies, saliva was obtained from 10 volunteers. Each subject was asked to provide two samples, one obtained by spitting directly into the Oragene collection pots, the other one using the sponge swab provided in the Oragene - DNA collection kit. A series of preliminary experiments were performed to optimize the extraction methods, ultimately to maximize the quality and yield of genomic DNA. **Table B1** shows DNA yield from the saliva spit (SP) and saliva sponge (SS) samples obtained from the 10 volunteers.

The average total DNA yield from the undiluted samples was 30.53 μg (range 9μg-110μg) and 22.44μg (range 2μg-57μg) for the saliva spit and saliva swab, respectively. The manufacturer (DNA Genotek Inc, Canada) recommends a minimum of 2 ml of saliva for an adequate yield of DNA. Each storage vial contains 1.9 ml of a DNA stabilization solution, equating to a total volume of 3.9 ml. Overall, the DNA yield increased proportionally to the saliva volume, **Table B1**; with DNA purified from saliva spit samples having a higher yield relative to saliva swab samples. DNA yield may be reduced by the presence of contaminants like glycoprotein (mucin), protein, salt, and alcohol during DNA collection and isolation procedures. As seen in **Table B1**, acceptable A_{260}/A_{280} ratios (1.8–2.0) were obtained from all but two samples (SP6 and SP7) demonstrating low protein and organic contamination of their DNA products. Furthermore, protein contamination in the saliva swab was far less than in saliva spit samples, the average A_{260}/A_{280} ratios being 1.82 ± 0.12 and 1.64 ± 0.12 (mean ± SD) for SP and SS samples, respectively. In addition, the A_{260}/A_{230} ratio provides qualitative information on the degree of alcohol, glycoprotein and salt contamination and thus provides information on DNA purity. The ratios of A_{260}/A_{280} and A_{260}/A_{230} should be in the range of 1.8-2.0 to ensure low contamination levels, with lower values reflecting impurities due to EDTA, glycoprotein, buffer, salts and ethanol that absorb at 230nm and protein and organic contaminants, which have absorbance maxima near 280 nm. On

this basis, the saliva swab samples provided lower contaminant values (**Table B1**) and therefore provided a better quality of DNA. These observations were supported by the similar values obtained from Genotek (Birnboim 2004) and other published data (Barker et al. 2004; Hansen et al. 2007; Rogers et al. 2007; Rylander-Rudqvist et al. 2006).

To identify the potential contaminants, genomic DNA from samples SP1, SP4, SP9 and, SP10 were precipitated again using 3M sodium acetate and ethanol to reduce the impurities. In this secondary analysis the DNA pellet was left to dry for 1 hour to ensure complete removal of ethanol. The outcome shows that there were only minor changes in both the A_{260}/A_{280} and A_{260}/A_{230} ratios compared to previous results (**Table B2**). Consequently, contamination at 230 nm by ethanol or salt carryover can be excluded as a significant issue in these samples. Thus, the observed 230 nm contamination probably reflects residual glycoproteins within the saliva samples.

To investigate the likely impact of increased hydration buffer volume or extending the DNA hydration period on final yield samples, SP5 and SP6 underwent both an overnight hydration step in TE buffer (50 [sample 'a'] and 100 µl [sample b]), as opposed to the recommended protocol of 2 hours at 50°C. This is an important step because if the DNA yield is low, one might consider using a reduced volume of buffer for resuspension; yet high yields of genomic DNA require a sufficient volume and time to be efficiently dissolved. The comparison of the 'a' and 'b' samples for SP5 and 6 illustrated no consistent difference between the yields of DNA with increased TE buffer volume (**Table B1)**. Similarly, the yields observed with the 2 hour and overnight incubation were broadly similar, samples SP5/6, 'a' versus 'c' and 'b' versus 'd' **(Table B1)**.

An overnight incubation step at 50°C has also been recommended by the manufacturer instead of standard 2 hours incubation to promote complete cell lysis and release of DNA from the Oragene - DNA/saliva samples. To test whether an increased incubation period enhanced DNA yields from the saliva samples SP7 and SP8 were divided into two samples 'a' and 'b' with the former undergoing a 2 hour incubation

and the later the overnight. The outcome was inconsistent between the two samples suggesting little additional advantage of performing the extended incubation **(Table B1)**.

Table B1 DNA yield and quality in saliva spit and saliva swab samples quantified by UV spectrophotometer

Sample	Total [DNA] Undiluted (µg)	Total [DNA] Diluted (µg)	Ratio A260/A280	Ratio A260/A230
SP 1	79.82	82.20	1.70	0.78
SP 2	44.59	44.05	1.72	0.93
SP 3	9.60	12.36	1.65	0.60
SP 4	80.43	78.30	1.70	0.92
SP 5a	9.34	9.15	1.69	0.83
SP 5b	14.27	12.56	1.68	0.82
SP 5c	14.29	9.13	1.70	0.87
SP 5d	15.46	12.71	1.68	0.83
SP 6a	18.24	16.65	1.50	0.59
SP 6b	17.80	18.90	1.51	0.57
SP 6c	20.94	22.45	1.51	0.60
SP 6d	19.76	22.70	1.52	0.59
SP 7a	13.21	12.72	1.47	0.51
SP 7b	8.93	8.36	1.38	0.43
SP 8a	17.31	20.20	1.80	1.16
SP 8b	23.22	23.40	1.76	1.04
SP 9	109.67	110.08	1.76	1.06
SP 10	32.73	35.00	1.73	0.96
SS 1	15.85	16.88	1.90	1.52
SS 2	2.20	2.60	1.69	0.99
SS 3	9.13	22.73	2.00	0.50
SS 4	40.05	41.38	1.66	0.51
SS 5	57.15	73.20	1.90	1.36
SS 6	56.83	70.35	1.86	1.23
SS 7	3.02	3.60	1.78	0.94
SS 8	2.83	3.41	1.65	0.71
SS 9	27.15	27.30	1.94	1.60
SS 10	10.20	11.25	1.85	1.11
Mean ± SD (SP)	**30.53 ± 29.25**	**30.61 ± 29.49**	**1.64 ± 0.12**	**0.78 ± 0.21**
Mean ± SD (SS)	**22.44 ± 21.76**	**27.27 ± 26.43**	**1.82 ± 0.12**	**1.05 ± 0.39**

Notes: SP=saliva spit and SS=saliva swab samples. For SP5 and SP6 'a' indicates that the sample was re-suspended in 50 µl of TE buffer and 'b' indicates that the sample was re-suspended in 100 µl of the buffer, with samples incubated for the standard 2 hours at 50°C. SP5 and 6 samples labeled 'c' and 'd' reflect DNA concentrations in samples re-suspended in 50 and 100 µl TE buffer respectively, but with an overnight incubation at room temperature. Samples SP7 and 8 underwent either the standard 2 hour pre-incubation, 'a' to permit DNA extraction, or an overnight treatment, 'b'.

Table B2 Comparison of DNA quality after genomics DNA clean-up

Sample	Total [DNA] precipitated* (µg)	Total [DNA] recovered* (µg)	New A260/A280	Previous A260/A280	New A260/A230	Previous A260/A230
SP1 undiluted	19.96	22.06	1.69	1.7	0.76	0.78
SP1 diluted		22.61	1.74		0.74	
SP4 undiluted	20.11	23.39	1.7	1.7	1.11	0.92
SP4 diluted		24.54	1.76		0.89	
SP9 undiluted	15.67	17.74	1.77	1.76	1.11	1.06
SP9 diluted		17.16	1.66		0.91	
SP10 undiluted	8.18	9.31	1.72	1.73	0.98	0.96
SP10 diluted		9.36	1.79		1.02	

Despite the protein contamination being within an acceptable range, quantification by absorbance is still subject to interferences that can impact on the absolute quantification of the sample DNA content. Therefore, an alternative method using fluorescence was employed to derive a more accurate quantification of the genomic DNA concentration. The picogreen (fluorescence dye) method uses a fluorophore that exclusively intercalates into double-stranded DNA, yielding values based only on DNA and avoiding the interference from contaminants. **Table B3** shows the summary of the picogreen results with comparison to the quantification obtained by absorbance. Overall, the DNA estimates based on absorbance were 3.5 fold higher compared with the picogreen results in both saliva spit and saliva swab samples. As shown in **Table B3**, there was a statistically significant difference between the DNA yield quantified using the two methods (Wilcoxon Mann Whitney test) ($p = 0.001$). Furthermore, the respective concentrations obtained using the two quantitation methods were broadly in agreement with previous reports (Mulot et al. 2005; Nishita et al. 2009; Rylander-Rudqvist et al. 2006), suggesting that quantitation by fluorescence, particularly picogreen, is the more precise method (Rengarajan et al. 2002). Despite the reduced concentrations determined using the Nanodrop, the concentrations using the two methods were highly correlated, as would be expected ($r=0.876$, $p<0.01$).

Table B3 Comparison of DNA yield by quantitation method

Donor	Picogreen Total [DNA] (µg)	Absorbance Total [DNA] (µg)	Ratio [DNA] Picogreen/ Absorbance	Saliva volume (ml)
SP 1	14.59	79.82	1.80	5.35
SP 2	13.83	44.59	3.10	5.50
SP 3	2.58	9.60	2.70	2.90
SP 4	18.26	80.43	2.30	3.20
SP 5	17.37	53.36	3.70	4.00
SP 6	26.72	76.75	3.90	6.40
SP 7	5.77	62.66	2.30	3.20
SP 8	19.29	40.52	3.30	3.20
SP 9	32.12	109.67	2.90	5.20
SP 10	14.95	32.73	4.60	3.20
SS 1	2.20	15.85	1.40	2.20
SS 2	0.20	2.20	0.90	1.95
SS 3	6.03	9.13	6.60	≈1.9
SS 4	9.74	40.05	2.40	1.95
SS 5	17.23	57.15	3.00	2.18
SS 6	17.66	56.83	3.10	2.22
SS 7	0.47	3.02	1.60	≈1.9
SS 8	0.79	2.83	2.80	≈1.9
SS 9	7.35	27.15	2.70	2.40
SS 10	3.84	10.20	3.80	2.10
Mean ± SD (SP)	**16.55 ± 8.71**	**59.01 ± 28.87**		
Mean ± SD (SS)	**6.55 ± 6.54**	**22.44 ± 21.76**		

The quality of the DNA extracted by using the Oragene DNA kit was further examined by agarose gel electrophoresis (**Figure B1**). A visible band of intact high molecular weight DNA (> 23kb) with limited or no degradation was observed indicating that the integrity and quality of the DNA was excellent. For lane 5 and 10, more than 100ng of total DNA was loaded. Conversely, 10ng was loaded in lanes 22, 23, 28 and 29. Furthermore, the absence of bands at the bottom of the lane indicates that the genomic DNA was free from RNA contaminants.

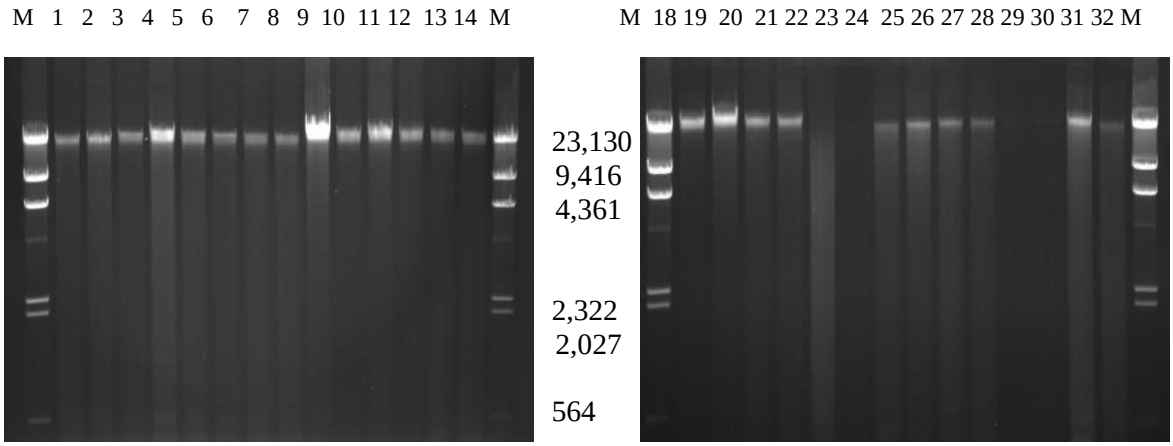


Figure B1 Integrity of DNA in saliva spit and saliva swab samples. Lanes 2-15 and 18-21, show DNA purified from saliva spit samples and lanes 22 - 31, DNA purified from saliva swab samples. A Lamda Hind III digest marker (M) was ran in lanes 1, 16, 17 and 32.

The objective of the preliminary work described in this section was to compare the quality and quantity of DNA extracted from two different collection methods; saliva spit and saliva swab. The Oragene DNA Saliva Collection Kit (SCK), employing a saliva swab appeared ideal for the collection, stabilization and purification of samples of young children. Although, DNA purified from saliva showed glycoprotein contamination, it has been stated by the Oragene technical support that its performance in genotyping assays is similar to the DNA purified from other matrices. The SCK provided DNA yields in spit samples for all participants greatly exceeded our baseline requirements for genotyping 96 SNPs on Illumina platform, with minimum of 250ng of total DNA. The saliva swab samples also provided sufficient DNA in all individuals except for few cases. Therefore the use of the SCK, coupled with RNA digestion, provide high quality DNA for genetic analyses with sufficient quantities for archiving for future studies.

Based on the preliminary results, an optimized extraction protocol was developed for the 'Schools' study samples, including a 20 minutes drying step (post-precipitation) to ensure complete removal of ethanol. Taking into consideration

sufficient buffer to dissolve the DNA pellet, 100 µl of TE buffer is to be added to the pellet followed by 2 hour incubation at 50°C. DNA concentrations were quantified with both methods to accurately estimate both the yield in the samples, and to provide absorbance information for quality control purposes. As many of the saliva samples of the school study failed to yield sufficient DNA yields for our downstream application, whole genome amplification was applied to all the samples, regardless of the yield, as described previously (Beckett et al. 2008; Chartier and Pinard 2005; Rylander-Rudqvist et al. 2006).

Appendix C

Assessment of DNA Concentration and Quality for Saliva Samples

Overall, the DNA yield increased proportionally to the saliva volumes **(Table C1)**. Therefore, DNA purified from Year 2 and Year 3 samples have higher DNA yield compared to Year 1. The average DNA yield recovered was 7.54 ± 2.96 ng/µl (1.79-30.01 ng/µl), 21.96 ± 14.71 (2.37-89.50 ng/µl), and 68.52 ± 54.77 (4.37- 459.44 ng/µl) for year 1, year 2 and year 3 samples, respectively.

Table C1 DNA Yield in Year 1, Year 2 and Year 3 quantified by Picogreen Assay and Absorbance

Description	[DNA] (ng/µl) (Mean ± SD)	Total [DNA] (µg) (Mean ± SD)	₸Saliva volumes (ml) (Mean ± SD)	A_{260}/A_{280} ratio	A_{260}/A_{230} ratio
Year 1 (n=173)	7.54 ± 2.96	1.53 ± 0.75	1.91 ± 0.20	1.83 ± 0.21	0.76 ± 0.27
Year 2 (n=413)	22.04 ± 14.76	5.96 ± 4.35	2.26 ± 0.34	1.85 ± 0.13	1.09 ± 0.34
Year 3 (n=402)	68.52 ± 54.77	9.86 ± 8.36	2.82 ± 0.64	1.82 ± 0.08*	1.13 ± 0.23*

Notes: * Absorbance measurement for Year 3 samples was only performed on 36 samples to minimize the usage of genomic DNA for future application. ₸ Include 1.9ml of preservation solution.

In addition, average $OD_{260}/_{280}$ ratios for Year 1 and Year 2 samples are within the acceptable range (1.8–2.0). This demonstrated low protein and organic contamination of their DNA products. Meanwhile, the $A_{260}/_{230}$ contamination may be due to glycoprotien (mucin) in the present work as DNA were extracted from saliva samples. In addition, extracted genomic DNA was still intact and in the high molecular weight (>20kb) range reflecting good quality and integrity of the DNA (see **Figure C1**).

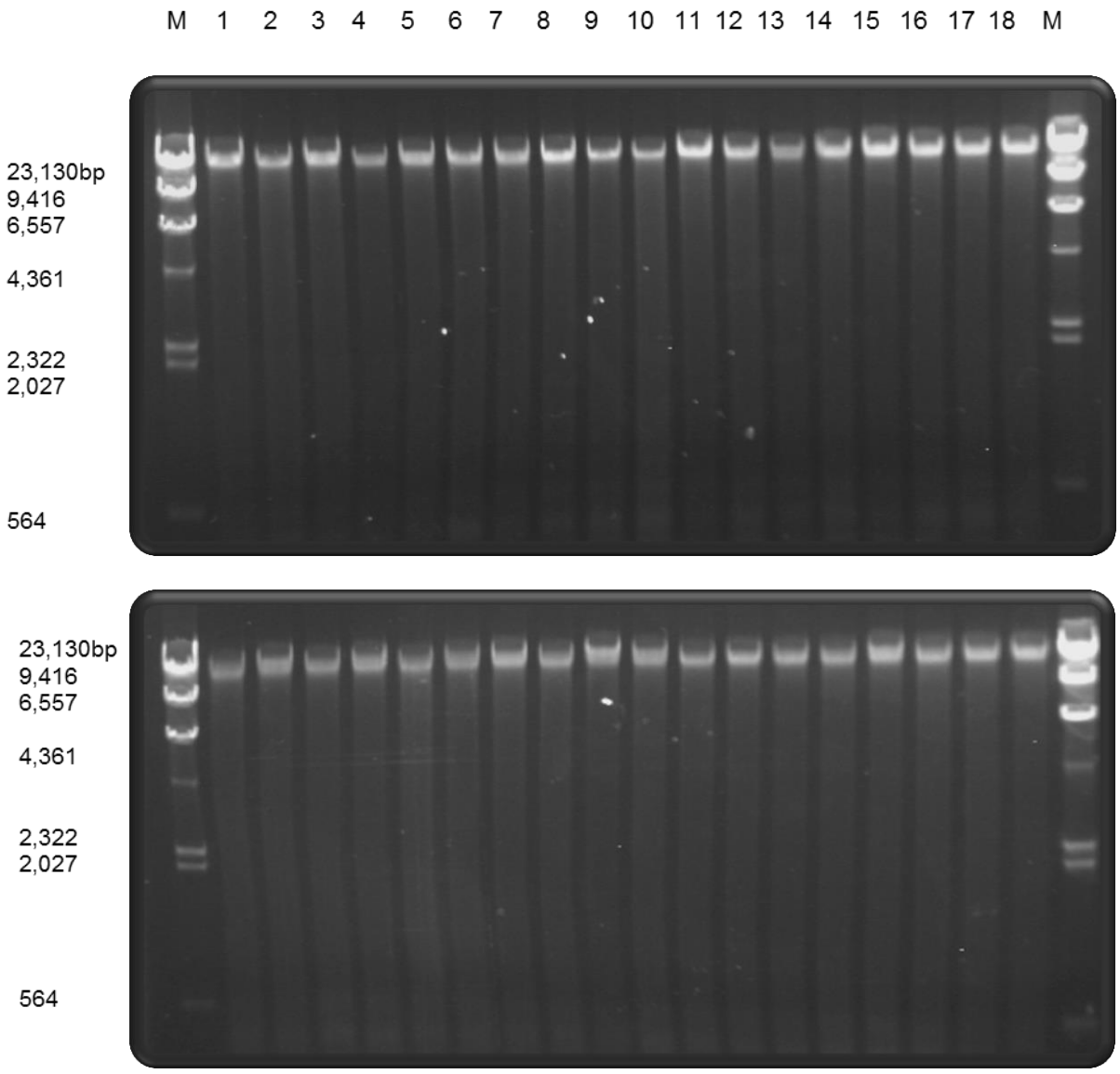


Figure C1 Integrity of genomic DNA purified from saliva samples. Approximately 100ng of gDNA from 36 samples is visualized on a 0.75% agarose gel (1 x TBE buffer, 90 V, 60 minutes) stained with ethidium bromide. Lanes 2-19: genomic DNA; Lanes 1 & 20: Lamda Hind III digest marker was used as the marker (M).

Appendix D

Genotyping Technology

SNP genotyping strategies typically involve two steps: allele-discrimination and allele-detection. Allele-discrimination is achieved by primer extension, hybridization, ligation, or enzymatic cleavage, whereas most allele-detection methods are based on mass, fluorescence, or chemiluminescence. The more straightforward and less costly options permit only a limited numbers of SNPs to be identified, whilst newer methods permit up to one-million SNPs to be investigated simultaneously, albeit at a greater cost and ultimately at a greater ultimate statistical penalty. Generally, SNPs analysis can be categorized in 4 groups based on their molecular method: allelic specific hybridization, primer extension, allele-specific oligonucleotide ligation and enzymatic cleavage.

Allele discrimination is performed using allele-specific biochemical reactions, with four main methods employed: hybridization, primer extension, ligation, and enzymatic cleavage. Allele specific oligonucleotide hybridization exploits differences in the thermal stability of double-stranded DNA to distinguish between perfect and mismatched target-probe pairs to achieve allelic discrimination (**Figure D1**). Allele specific probes are then immobilized on a solid phase, 'target SNPs', to which fluorescent labeled PCR fragments are hybridized. Subsequent to stringent hybridization and washing, the unbound targets are removed. From the location of the probes, one can determine the genotype for the DNA sample. As hybridization does not involve enzymes in allelic discrimination, this is the easiest method of SNP detection. High-throughput SNP analysis such as Affymetrix and Illumina platforms employ this approach. Consequently, the accuracy of the probes determines the success of the method.

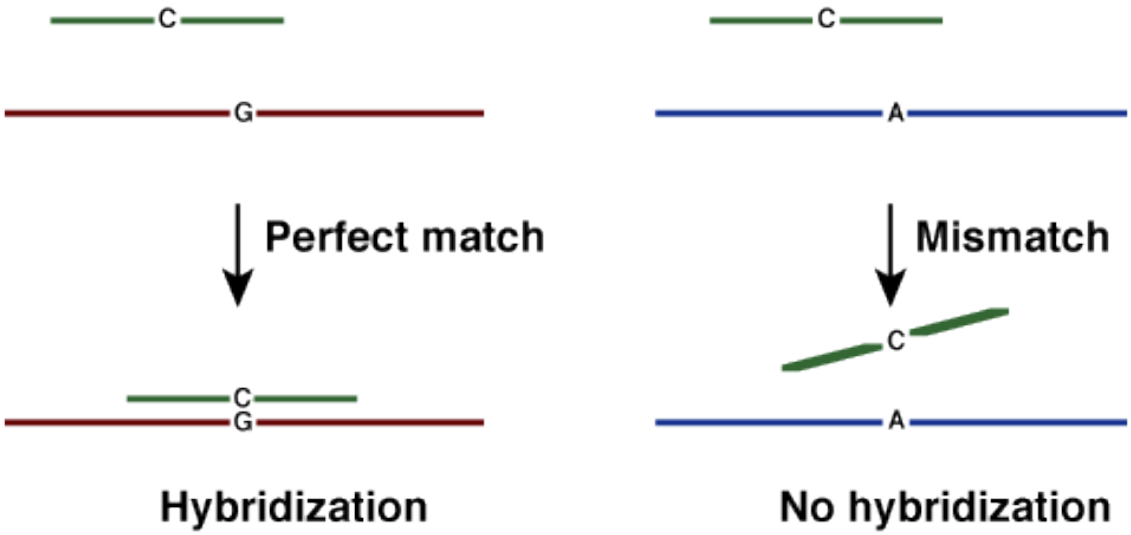


Figure D1 Allele-specific hybridization. In this approach, two allele-specific probes are designed to distinguish between two alleles on the target DNA sequence with one mismatch base. Hybridization only occurs with a perfect match to the DNA sequence. Source: Kwok (2001).

There are two methods of SNPs detection by primer extension: sequencing (allele-specific nucleotide incorporation) and allele specific primer extensions. In both approaches, PCR amplification is performed, followed by single base extension using primers that anneal to one base of polymorphic site. Sequencing determines the alleles (homozygote, heterozygote or wild type) present in the DNA samples (**Figure D2**), with the sequence of amplified DNA subsequently identified by mass spectrometry or fluorescently labeled primers, where the PCR products are run on gel electrophoresis or capillary electrophoresis. Recently, the sequencing field has progressed rapidly with the production of instruments capable of typing millions of DNA sequences in a single run making GWA studies rapid and cost effective. The Illumina genome analyzer, Applied Biosystems SOLiD sequencer and, Solexa genome analyzer offer this approach.

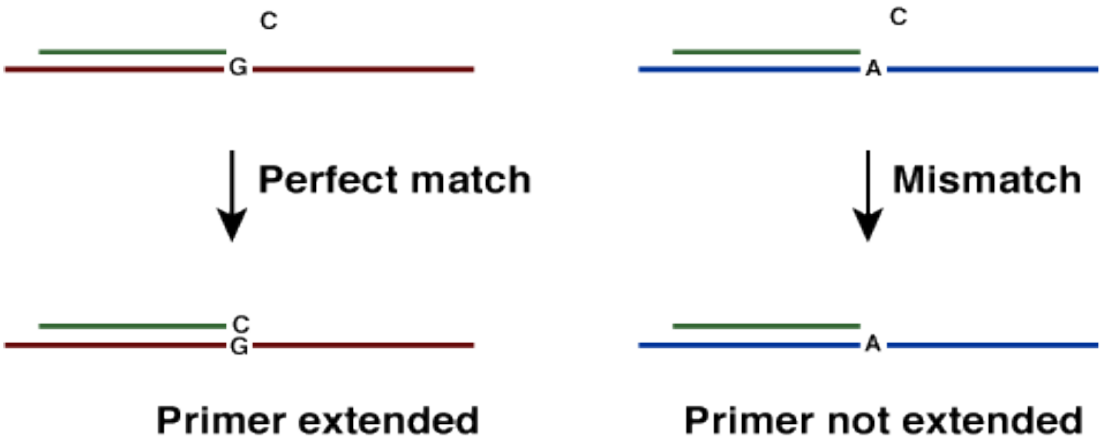


Figure D2 Allele-specific nucleotide incorporation. Amplified PCR product is sequenced to determine the presence of the base at the locus. This could be achieved either by fluorescence labeled primer or mass spectrometry. Source: Kwok (2001).

Alternatively, in the allele specific primer extension method, the target sequence is only amplified when the 3' end is complementary with the template after successful extension of the one base of the polymorphic site. Otherwise, the primer extension or amplification of the target sequence will be terminated (**Figure D3**). Therefore, by knowing whether the PCR product is produced or not, one can determine the allele types by running the PCR products on an electrophoresis gel. One method which employs this approach is PCR-DGGE (PCR-denaturing gradient gel electrophoresis), which employs fluorescent labels in the SNPs detection. This approach is robust and minimal primers are required to carry out the genotyping. In addition, primer design and optimization is relatively straightforward.

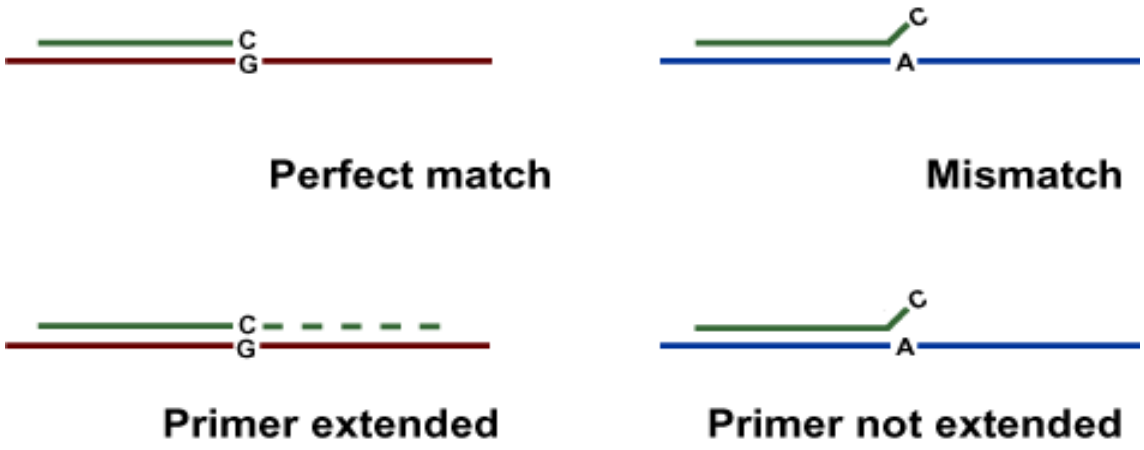


Figure D3 Allele-specific primer extensions. Primer is only extended when it is complimentary to the base at the locus. Therefore, the allele could be inferred only if PCR product is generated. Source: Kwok (2001).

In this allele-specific oligonucleotide ligation, complimentary oligonucleotides with alleles specific to the polymorphic site are designed at its 5' and 3' end. By using DNA ligation, two adjacent oligonucleotides are joined to each other after annealing to a DNA template (**Figure D4**). Ligation can only occur when it perfectly matches the sequence of DNA on the template. This method is very slow in the detection and requires a high number of the probes in the analysis. However, ligation is the simplest way of optimization and very accurate.

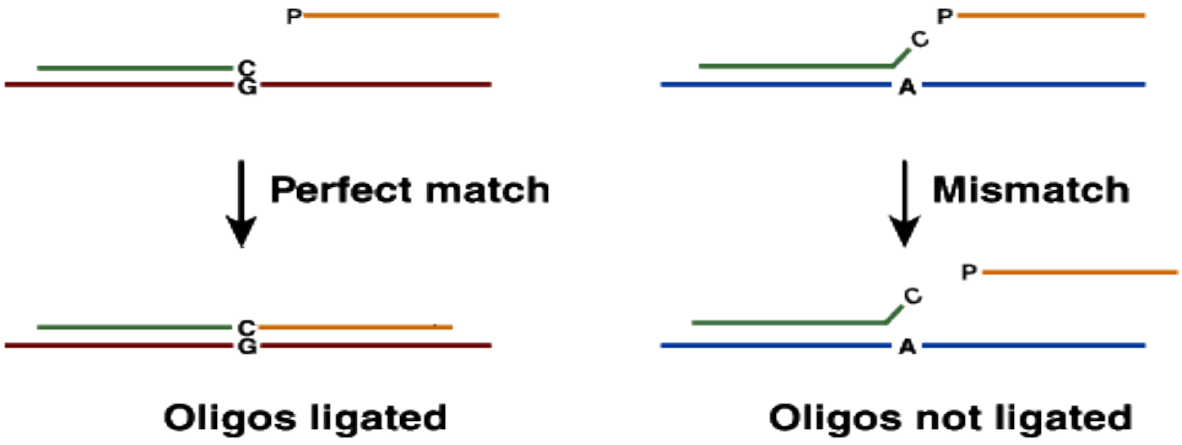


Figure D4 Allele-specific oligonucleotide ligation. The presence of allele at the locus is determined only when oligos are ligated with probes. Source: Kwok (2001).

Enzymatic cleavage for allele discrimination is based on the ability of certain classes of enzymes to cleave DNA by recognition of specific sequences and structures. Such enzymes can be used for discrimination between alleles when SNP sites are located in an enzyme recognition sequence and allelic differences affect recognition. This allelic discrimination method employs 5' – 3' nuclease activity with two probes (reporter and universal) for each allele, such that the SNP located at the overlapping site of the DNA strand should be centered over the mismatched base where the probes only differ by one base.

The probes for these assays have a lower T_m requirement than the DNA polymerase in PCR assays. During the PCR amplification, once the specific enzymes (probes) have bound to its specific piece of the template DNA after denaturation (high temperature) and the reaction cools, the primers anneal to the DNA. Subsequently, *Taq* polymerase adds nucleotides and removes the probes from the template DNA. They are dependent on each other because both probes and primers are working simultaneously and therefore, must hybridize to their targets for amplification and cleavage to occur. When labeled reporter is separated from the cleavage site, this allows the reporter to

emit its light. The more amplification takes place, the more chance for the reporter probe to bind and in turn more emitted light is detected (**Figure D5**). Thus, SNP determination relies on the signals generated if the target DNA sequence is successfully amplified during PCR. The TaqMan allelic discrimination assay and PCR-restriction fragments length polymorphism (PCR-RFLP) are an example of this approach.

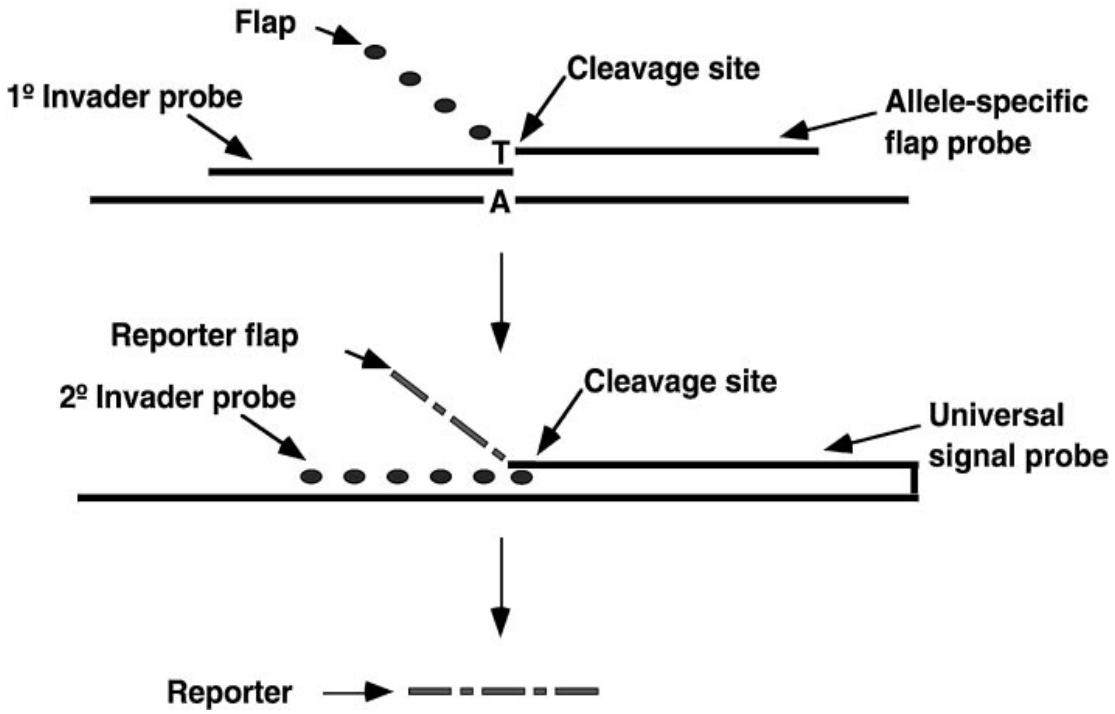


Figure D5: Invasive **cleavage allelic discrimination.** Probe will only bind to specific target but not with a one base-mismatch. The second probe is employed to identify the polymorphic site at the locus. Source: Kwok (2001).

Table D1 Single nucleotide polymorphism sequences

SNP Index	SNP ID	Base Change	Location	Top Genomic Sequence
1	rs2917666	[G/C]	Intergenic	ATTATGAAAGGGCTTATTGGCAGAGAAAAAGTAGAGGGGTCCAGAGACTCAAGTTCCTAA[C/G]GAGGCTGAAGTAAGAGTGGGGGATTAAATAAGGGCAAAAACTGAGAGAATAAAAAAAGTG
2	rs7782389	[T/C]	Intron	AAATGCACATTATTAATCAAAAAACAGCAAAGTACGTATATGTAAATTGTATCTATAGAC[G/A]GAGGGACACCAGTGGAGAGAGGTTTAGGAGATGGCCACCTAGTTTCCTTACATGTTTCCT
3	rs997279	[A/G]	Intron	GCTTGTGCTACCCAGAACACGGTATTATGCAGTTTAAGAGATTCTTTACATGCTTGGCGC[A/G]GTCAAACATCAATAGTTCTACCTTGATAGACTTTCAAGTCTGGAAAAAGCTTAATGAAGT
4	rs4147581	[G/C]	Intron	CCTACCTCGAACTGGGAAATAGACCACGGTGTAGGGCGGCACTGCAGCGAACAAAGAAAA[C/G]TTTGCTGGAGCCCGGGGGAGGATGCCAAGGCGCGGTGAGCGCAGCCTGCCCCCCCCTCCC
5	rs3759757	[G/C]	Intergenic	ACGTGGGCCAGACCAGCACCAATGGAGGACCACCAAACCCATGTACAATGAGGAGCTCTA[C/G]ACTAACGCCACTGATGGCGGCCGTCATCCATAAGATGCCCCTGGGCTATGACCACTTCAT
6	rs1137933	[A/G]	Nonsynonymous	ACACCTTCATCTGGCCAGCTGGGCTGGCTGGGTTACCTCCAGGATGTTGTAGCGCTGGAC[A/G]TCACAGAAGTCCCGGACTCCGATCTCTGTGCCCATGTACCAGCCATTGAAGGGGCACCCT
7	rs2826003	[A/G]	Intergenic	CCTAAGTGTTAATCATTTGTTGTTTAATGTGGTTTAGAAACTGTTCTGTGTGCTTTGCCC[G/A]TATTACTTAAATTCTGACAACAACTCTACAAGGTTGGTTCTATTTTCCCTTATAAGGGAT
8	rs17072738	[A/G]	Intergenic	GGGCCTTGGAAGTATTCTTAATTTTCATGAGTCAAAAGGAAGGAAAGGTAGCAAATACAT[A/G]TCAATGACTTAATTCTAATCATCTTCTGTCCTATATCTGCCCACTTCTCTTTAAAAAGCA
9	rs2246012	[A/G]	Intron	TGTTATTATCATTATTACTGTTTTATTTTATTTTATTTCACAAATAGGGGTCCTCTTGCT[G/A]GTAGCATACTTCCTGCTGCACCTTGCCTTCCATTTGCTGTTCCCCCGGCTCCCCTACCAT
10	rs2074113	[T/G]	Intergenic	ATATTACATACACTGTATGTTTTACATGCTAAGAGAGACTTACTGAAACATAATGTACAT[C/A]CAATAAAATAAAAGTGAAAAAGTCTCACTTACTTGGCATCACAACCAATAGGTGTGAAGT
11	rs7232792	[T/C]	Intergenic	GTCTGCAATGGAGAATCCACAGGTATGGCCAAGGTTGGTTCAGTTGAGTTCAGCAAAACT[G/A]ATGGAACACCAACTACATGGCACGTACTTTGCTAGGTACAATTTTATTAAAAAAATGAGT
12	rs2284659	[T/G]	Intergenic	AGTTTTCTGAAGTTGCAGCTGAACACTTTATGAAGCACTTTCTCATATAGTATCGTATTT[C/A]ATGCTTGCAATACTTTTTGTGGGGTGGATATTAATTTGGCAACTTTGTTTATAAGAAAAT
13	rs2781668	[A/G]	Intergenic	ATGATATAAAATGTCCTAGTAACAAAAACAGGATGCCCTTAACACTGTTGGGGACAGGGT[G/A]TGTCAGTAAAGGCTGCAAGAAATAAATGCTGTTTAAGCTGAGTCTTGCGACATGAGTAGG
14	rs2066853	[A/G]	Nonsynonymous	GTAAAAACAGTGACTTGTACAGCATAATGAAAAACCTAGGCATTGATTTTGAAGACATCA[A/G]ACACATGCAGAATGAAAAATTTTTCAGAAATGATTTTTCTGGTGAGGTTGACTTCAGAGA
15	rs2733262	[T/C]	Intergenic	GAGGCATCAGGAGATATTCACAGCCAGAATAGAACAGATGAAAATCCAGCATTACCTTTA[G/A]ATACCACTCTATACAGACGGCTCGCTCTGTTAGGAAGAGGTAAATAACATTTCAATTACA
16	rs1138272	[T/C]	Nonsynonymous	AGCAGGGTCTCAAAAGGCTTCAGTTGCCCGGGCAGTGCCTTCACATAGTCATCCTTGCCC[G/A]CCTCCTGCCAGACACCACCATGTATCATCCTACTCTCATCCCTCCCAGCCCCTGAGACAC
17	rs17599586	[T/C]	Intron	ACCGATGGAAAACACTTTAGTTATTGCATCATGTGACTGATGGACATTCTGAAAGTGTCA[G/A]CTGGGACTTGGTTTCAGGTCAAGGAGGTATAAATATATATTAGTCAAGTGTACACTTGCT
18	rs742869	[A/G]	Intron	TGTGTTGCTCTAAACTTACCCTGTGGGTGGCAGAAGAAAGCAGATTGCATACCCACGGAC[A/G]CACATGATAATGATCTTGGTGAGAACAAGAATTCTTGATTAATCCTGTCCATTTCTCCCT
19	rs17883901	[A/G]	Intergenic	GGCCTGAAGCCCGGTGGGAACCGCGGGGGCTCAGGAGGCGTGTGCAAGGGTGATTGGGTC[A/G]CAGTTGAGGGGAGCAGCTCGCGAGAAGGGGGCGTGCGGTCGGCCTCGCTTTACGCAAACG
20	rs2364723	[C/G]	Intron	TCTTAGTGCCAAATTCTTCCTGCATACTGAAAAGTTAACCCAGGCTTGAGGAACAGTTAA[C/G]GAGGAAATATTGACACAAAACCCCCTTTAAATGTCTTCAGTCAGGACAGAGGAAATTCCC
21	rs8078340	[T/C]	Intergenic	TAATTATAGTACAGAGTGATATCTGTAGTGATGGAGGGATGGTATGGTGCTGATGTAGCT[G/A]GTTCTGCCTTGTTCAAGGAAAGTTTACAAGGGAGAGGACACCAGGGCAGGGTTTTGAGGG
22	rs3742879	[A/G]	Intron	TAAATATGCTAGAGTCTCTTGCCTGCAAAGGATCTTTTCTGCTATTCCACATCATTGCAG[A/G]GTTTGTTTTTGCTTAAACTTTTGTAGCTAAAAATTATATGGCCATGCTAATAAAAAGGAT
23	rs268691	[A/G]	Intron	GAAGCGAGAGAACTAGGGTCAAGGATTCTTGGGACTCCAAGACGCACTGGCGGGGGAAGG[G/A]GGTTGCCGGGAAGAGTTCAATGTTCGGGTCTTGGAGCCGGACAGCGAGTTTGGAGGTGAA
24	rs2301022	[T/C]	Intron	TTGATTTCTTGATGCTCAGAGTCACACACCACAGTTTGTAAAACATTGTTCAAAGGACTA[G/A]TTCTGGTTTCTAGTGTTTTATTTAGGGTGCTTCATAACAGTAGGATAAAACTAAAGAAAC
25	rs3170633	[A/G]	Intergenic	TATGCTCCTTTAACACTGTCTTTCTAGATTTTTCACCCAGTATTTTCAAAATTTGGGAAT[G/A]TAAACAATTGATATATTTATTGTATGTTGGCTAGCAGTTCATCCTTCTGCAAAATATGCA
26	rs2779251	[A/G]	Intergenic	CCTACGGCTTCCTGGGCCTGCACGGCTGCCTTCTTCTTGGCCCTGCCCACCCAGCAGGCC[A/G]CATTCCTGGCCTCTTCTGATGATGGATGCAGCAGGTGAATTCTGCTGGGGAGAAATCAAA
27	rs7132743	[A/G]	Intergenic	CTTCTCCAAAATGCTTTTCTTGATACGTTTAGTTGGAGTCCCTTTTTCTTCTTTGTTTTC[G/A]TGGTGCTGGATTTATATCCATATTTTTCCTTTTACAATACTTATTTCACTAACTTATTCA
28	rs1871042	[T/C]	Intron	ATGGATCAGCAGCAAGTCCAGCAGGTTGTAGTCAGCGAAGGAGATCTGCAGGGCAGGGGG[G/A]CCAGGGACTCAGCCAGCGGGAGCAGGCTGGAGGAATTGGGGGCAGAGGGCCTAAGCCCAT
29	rs6770096	[T/C]	Intergenic	TGGTTTCAATTAACAGATAATCCTTTACAATAAATAATTGAGATCTGGTTTCATCACTTC[G/A]CTAATTGTTTCTTCATACATCTGGCTCAATGGAATATACTTTGCCTGAGATATGTTATTA
30	rs7557529	[T/C]	Intergenic	AGGTTTTTTTTTAATTCTAATATTACTCAAAGGGCAAAATGAGAGTCCATGGAATAGAAA[G/A]TGTCAGGAGACAATTATGTTGGCAAATCTCATTTCCCTTTTTTATTTTCCAGCCTCCAAG
31	rs2779253	[A/C]	Intergenic	AGGCTGCAGGTTGCTATGTCTCTACTTGATGCAGGTCCTCAGGGAAGCAGGCTGTCCCAC[C/A]AAAATCAGAGCTCAGAACTCAGGGTCAGGTATCTACTCCCATATGtccatccatccatcc
32	rs2606345	[A/C]	Intron	TGTGACTGGTGAGTTTTTAGGGACTGGTATTTCCAGCTCCCTGCAGTTGGCAATCTGTCA[A/C]CCTGATTGTCTCCCAGCAAAGGACAAAGAAGATGGTCAATGAGATAGGAAAGGCCTGGGA

SNP Index	SNP ID	Base Change	Location	Top Genomic Sequence
33	rs8192287	[T/G]	Intergenic	AGGCTGTCCACCCCGCACGCCCACCGTGAAGATCCAATGAGATCATTACTCGCCCAGTGA[C/A]AACACAGTGGCTGGCACTAGCCGCACTCATAAGTAATCATCATTGTCATTGTCGTTACCG
34	rs2781666	[T/G]	Intergenic	ATTTTATGTGAAGTCTCATCATTGGTGCCATTATACTGATGAATAAACTGAAGCTCAGAC[C/A]CGTTGCCAACATAGGCTGTAGAAGAAAAGTCGGGATTTAGTCTCTGCCATTCTGAATCCA
35	rs4795080	[A/T]	Intergenic	GGTGAGAGAGGGCTGCCTTGAGCCTCAGACCGGTGGTGTCCCTAAGATGCTCCTCCCAGA[A/T]GCATTTGGAATCAAATGGACCTGAATCTTGGTTCTGCAGCTTTAGCTGTATGATCTTGGG
36	rs7144243	[T/C]	Intergenic	ACGCCCGGCCTGGACGTGAGTATTTTTAAAAAGCTCCCAAGTTATTTTTAGGGAGCTGCC[A/G]GAGTTGAGCACTCTTGGGAGGATATCTTTCCCTCTGCCATCCCAGGCTCTAAAGTTAGAG
37	rs17722841	[T/C]	Intron	AAAAAATATTATTGTCAATTTCATAGGCAAAAGTGTGGTTGGTAATTAAAGTTATATGTT[A/G]GGATCATGGTTATGTATAGGAGAAGCACTTGTATTTGCAGATTTTCTACTTAGGGTGTCC
38	rs699473	[A/G]	Intergenic	TAAGTCAATAACTTATTGACTTATTTTTGCTGCATGGAAATGGGCACCTTGCAAATGTGC[G/A]TGGCATTTGCTCCAGAGCTGTGCTGTCATACCTATGAGAGCAGAGCAGCGGGTGGGTCGT
39	rs1889022	[T/C]	Intergenic	ACTTAATTTATGGAGTGCTCAGCTGAGGTTGCCACTAGCTGCTTACGAGAGGCCAGCAGC[A/G]CCGTGAGCCAGGCCCACTTAGGGAAGATTCTTGGCTCATCCATGTGACTCACCTTCCTCA
40	rs9282799	[T/C]	Intergenic	ACAACTACATTAGGGAGAAGTTGAGAAGCAAAGGGAAAAAAAGTCACCCTTGATCTCACC[G/A]TCCCAACACTTTGCTACCACTGTCCACTCATTTCCTTCTTGTCTTTCCCAAGTTAAATTC
41	rs7591449	[T/C]	Intron	GGAGACTGAGGCTCAGAAAATGGGGAGATTTGTAGCCAGGTCATTTGGCTTTGTAGGATA[G/A]GAATTCAAGTCCACGCATGCCAAATCTAAGCTGAGAGGTTAATTTTCAAAATCCACCACC
42	rs3894194	[T/C]	Nonsynonymous	CAGAGACAATGACCATGTTTGAAAATGTCACCCGGGCCCTGGCCAGACAGCTAAACCCTC[A/G]AGGGGACCTGACACCACTTGACAGCCTCATCGACTTCAAGCGCTTCCATCCCTTCTGCCT
43	rs1442293	[C/G]	Intron	CCCCTCCTTGGGGCCTCTGAGTTCCTCAGGAAGTATTTCCAGTATGGGCTGTGGGTTGCT[C/G]CAAGGCCTGGGGTTTTCTTGGTAATAGGCACTTTGCCACATGAGCATTTGACAGTCCAGT
44	rs749174	[A/G]	Intron	GGGTAAGGAGATAGAGATGGGCGGGCAGTAGGCCCAGGTCCCGAAGGCCTTGAACCCACT[G/A]GTTTGGAGTCTCCTAAGGGCAATGGGGGCCATTGAGAAGTCTGAACAGGGCTGTGTCTGA
45	rs689453	[T/C]	Synonymous	AAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGGA[G/A]GCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCC
46	rs3796644	[A/G]	Intron	GCTTTATTTTATTTTAACTCAGTCCCAATGGTTAATTTAAATCCTTTCTTTCTTTCATCA[A/G]TGATACAAAGAGCTCACTTTACTACAGTCTCATCTTGACAGAGTCAAATGCAGAGGCTGT
47	rs2234922	[T/C]	Nonsynonymous	TCCCCCCAGGGCTGGACATCCACTTCATCCACGTGAAGCCCCCCCAGCTGCCCGCAGGCC[A/G]TACCCCGAAGCCCTTGCTGATGGTGCACGGCTGGCCCGGCTCTTTCTACGAGTTTTATAA
48	rs17779352	[A/G]	Synonymous	CAACTTATTAATAACATCTTGTGGGAAAGGCAGCAGGCTAGCCAAACGGTCCAACTCTGT[G/A]TTAAGTCGGTCTCTATGCCGCTTGGAAGGATTTGACTTGATTCCTTCAGCTGGGATTGGC
49	rs2128238	[A/T]	Intergenic	tagagttaacctgtcctttatgccttaaatcttggtgcttgctttttttcccatgctaat[A/T]aatgtccttttgccatttttccccccagaataggacactttcatctatattaaatacttg
50	rs1799814	[T/G]	Nonsynonymous	AGGTGTTAAGTGAGAAGGTGATTATCTTTGGCATGGGCAAGCGGAAGTGTATCGGTGAGA[C/A]CATTGCCCGCTGGGAGGTCTTTCTCTTCCTGGCTATCCTGCTGCAACGGGTGGAATTCAG
51	rs10454231	[A/G]	Intergenic	TACGTTATTAGAAACTCCTACGAAGTGAATCTAAATGCCACGGTCCTTGTAGGGGGCTAC[A/G]TGGCTGATGCGGCAGGTCAGGAAAATGCATACCTTTGCGTGAGTTCCTCCAAACTCAAGA
52	rs7776785	[T/G]	Intergenic	ccttgattcattcagtcattcaattacctcatctgtgaagtgagggTAGTGTGTAGGGAT[A/C]AACTGGAACACAGAGGGGGCTTTAGGAAGCCAAGAATCACATCAACTTGTATAGAAAATA
53	rs8192290	[T/C]	UTR	CGCAGTACTGAATTGCTGAATTTTTAATCTGTAACACAAAGGTTGCCAAACATTCCCCCA[G/A]AGGAGCAGCTCTCAGTGGGGTGGCGGTGGATGCGAAACAGCTGAAGACGCGGGGCAGTTG
54	rs7140310	[A/C]	Intron	ATCTGTGGTAATCACATTGATCAGTGGTCCCCAAACAGTGAGACTTGACCAGAACTCAGA[C/A]CAGCAGACAGACTTAATAAATTTCCAGAGACTACACTGAGAAAGGAGGCAGGAGACTTGC
55	rs1800566	[A/G]	Nonsynonymous	CTTCCAGGATTTGAATTCGGGCGTCTGCTGGAGTGTGCCCAATGCTATATGTCAGTTGAG[A/G]TTCTAAGACTTGGAAGCCACAGAAATGCAGAATGCCACTCTGAGGATACAGAAAGCACAG
56	rs17861115	[A/G]	Intergenic	AATGTCACGCACATGGCTTCTCTCACAGGAACTGTCCAGCCAACAACTTGACACCATCAT[G/A]GGGTCTTCTGGAGATACAATTTGCCTCCCACAGTGGTTCGGATGGTGGCCATTTAGGTAC
57	rs11735827	[T/C]	Intron	CTTGGGATCTGCATGACTACTAATGTCAGAAGCTGGTTACTATTTTAAATAGAACACCCA[A/G]TTTAGCACCAGAGAAGGTATTCACTAAGGAACCTCCAACCCCATCCTGCTGAAAAATGAA
58	rs2531866	[T/C]	Intergenic	AGAAGTAGGGCTAGAAATGTTGGAGATGGACTTGAAGATGAATTTATATGGAAGATAAGG[A/G]AGCAGAGGTGACAGCAGTGACTTCTGGGTTTCTGGCTGTTGAGACTGGGGGGCAGGGATG
59	rs9900426	[T/C]	Intergenic	tgagcaggcaccatccaattggctgtaggccccaagagaacaaaaacagaggaaaggcaa[A/G]ttggtttagctttcccctggagctaggatacacttctttctcctgtccttcaatatcaga
60	rs10517	[A/G]	UTR	TTGATAAAACTCTAGGTTTACAATTGTACCCCAAGGTCATGGGACTGGACCCCATCCCAT[A/G]GTAAGTCATCAGTTTAGCAATGATAAAGAAAATAACCTTCTGAAAATTTGTATAGATCAG
61	rs2282885	[A/G]	Intron	AAGCTTGCAGGGCTTCTAAACAGAAGCATGGCATGCAGTCAGCTGGAGTCTTGATTTCAT[A/G]ACAATTTTACAGATTGGGACAAGATTTTTGAGGTAAAGGTTTCAAAAAATAATAAGTTAC
62	rs2001350	[T/C]	Intron	CTTCCTGGGCTCAAGCAATCCTCTCACTTCAGTCTCCCTAATAGTTGGGACTACAGGCTC[G/A]TATCACCACGTGCAACTAATTTGTTTTGATTTTGGGTAGAAACAGGTGTCACTATGGTGC
63	rs527705	[T/G]	Intron	CAAAGAGAATGACCAAGGAGTTAAGAAAGCTACCTCCAGCCATGCTACTCAAGTATGCAA[A/C]CCCGTCAGTGGCTTTCAAGGACACCTGTAAGGTGCGGGGGGCGGGTAGGGAACATAGAGA
64	rs4902503	[T/C]	Intron	TTAGTCATTATGAAGATTCACATAAATTTGAAAAGCACACAAATATTATGAAATAGAAGG[A/G]AGGAAGGGAGGTCTCTTTCTTTATTTTCAACAGGGAGAATTAAGCCTCTTAATTTTTTTT

SNP Index	SNP ID	Base Change	Location	Top Genomic Sequence
65	rs7875663	[A/G]	Intergenic	TGAGAGGTGCCTAATGAACACCAGAGGAAGAGAATCTGAAAACAATTCCTAACAATGTGG[A/G]GTGTCGTTAGTTAAATCCAGACCTCTACTTCTGTGTCTGATTCTTCTTTAACATAGAGCA
66	rs407257	[G/C]	Intron	TGCCCCCTTTGCCAACCTGGTGGGCAGGTGCTGAGTTCACAAGGTCCTAGAATCCCACAA[C/G]GAAGCCAGGGTGCCTGGTGGGAGCCCAGGGAGTCCCAGCTGCTGTTCATTCCCCCTTCTC
67	rs2531872	[T/C]	Intergenic	TGTTTGGTGAACATATTATGCACACATTTTGTGCTATTGTCTTAGTAGTTACTGCCAAGC[G/A]GTTTCTCAAAGCAGTGGTACCAAAGCACACTCCACTAGTAAAGTTGCTTCACATTCTCAA
68	rs2472299	[T/C]	Intergenic	AAATGAGGTGAGAGAAGAGTAGGAACTGGCTCACATGTTGAGCTGAGGAGTTCAAGCTTT[A/G]TTCTGAATGCCTGAAGAAATTGATAAGGGATTTTGATTAGAGATCTGACATGCTCAGATT
69	rs221454	[T/C]	Intron	TTTTTTCAGATTTTAATCAATAGTTGATGTACAAGTGGTCAGAATAACTCAAATAGGGTG[G/A]GAGGATCAAATGAGGAACCTGTAACATACCTGCTCATTCACCTATtggaagaaagttctc
70	rs2781667	[A/G]	UTR	GTGTATTCAATGTTGCTGGCATGTGTGGATTCACAGCTCGTGGCTGTAAGGAAATCTGGG[G/A]GTTTCATAAATCCATTAAAATCCTTTTAGAGTTGCCAACCCGCAGTAGGATTTTCTTTCT
71	rs13306703	[T/C]	Intergenic	GCTAAAGCCTAGATTTACTTTCATTTTGATGGCATTCTAGAGGAAATCTGCCTTTTTCTC[G/A]CCACCGCCCCCAACCTCCCCACCATGGCAACAT
72	rs2198843	[G/C]	Intergenic	GTATTCCTGGGGACCTAGAAGCCAGGTGTCCTCCCCAAGCATCACAAGGGTTAGACCACT[C/G]TGCAGGGTATTTAGGAGTCCAGGGCACTTCTGTGTGTTACCCACAACCCAGATGGATGAA
73	rs778233	[T/C]	Intergenic	agattgctgtgtcatgcagtaattctatttttaatttgtttagaaaaatccacacaattt[A/G]cataatggctataccattccacatgcccacaaacactgtaaaagattccctttttctccac
74	rs7156352	[A/G]	Intron	TGGACACTTGGGTTGCTTCCACCTTTTGGCTATTGTTAATACTGCTGTGAACATGGGTAC[G/A]CAACTACTTCAAGATCCTGCTTTCAGTCCTTTTGGGTACATACCCAGAAGTGGAATTGCT
75	rs2074175	[T/C]	Intron	GTGCGGCTGCAGAGTCAGATGGGGCGGGGATTTCGGGGCACCGGGTCCTCACCTTCACGA[A/G]AAAGGCCCCACAGCACGTCCCCACTACCCGACGACTCACTCTTCGTGGCTTCTCTCTCCT
76	rs3824781	[T/C]	Intron	GGGGACTGGCTGTCCAATCTGCAGGGCTGGCTGCAAGGGCTGGAGCAGAAGTCAGGCCTC[G/A]GCTCAGGCCATCGGAGGCTTCTTCCTTCTCTTGCCCCTCCCTCTCCCAGGCCTAATCACA
77	rs2236687	[T/C]	Intergenic	AAAACTTTACAAGAATAAAGAAAAAAATTCCCAAAGTAGCAATTTTAGAAATTTAATCGG[G/A]AAGATTATCAAATAGCTGGTCTTTCAGACAGATGAAACATTAAGCACTTATACCTACACG
78	rs10459953	[G/C]	UTR	CAGTCCCCTCATCAAAGGTGGCCGAGAGATTTTAAAGCAGGAATGAGGCTGAGTTCTCTG[C/G]GGCCGGAGCCTCAGTTTTCGACTCGCTACAAAGTTATGAACACACTGGCAGCCAAGAAGT
79	rs10483801	[A/C]	Intron	GCTCTAACAGATACAGATAACACATTAGAGTTCAAAAGTGGGAATGCAGTTTGGCTAGTG[A/C]AAGTGGCAGATTTCCTCTGCCTTCATACTTGATTCATTTCCAGGCAGGCATCTGGAAGGT
80	rs2153747	[T/C]	Intergenic	AAAGAAGCTTGTCATTTTCTCTTGTGTCATGGTAAGTTGCTACCAAATAAGCTTTATATT[G/A]TTTGTGACCAAAAGATTTTCTTCTCCCTTTCCATGTTATTTGACAGAGTCATACTGACAA
81	rs10853181	[A/G]	Intergenic	CGTTATGTGGCAGTTACTGTCTAACTCAGGGAATCCTTCAGGAAGATCCTCAGCTCCCCT[A/G]CTACCTTCCACATACCCAAGCTGCTGTCTGTCTGAATATGCTTTTCCCCATGTAACTCAG
82	rs1014025	[T/C]	Intergenic	CTGCCACATACCAAACACTGTCACACTCTGTGGTTCTCCTCGAACAGAGAAACAAAGGAC[G/A]GAGGACTCTCCAGGCCTCTCAAAACAGAATTTGGACATCTTCAGAAATGCCCCCAGTGTG
83	rs2779248	[A/G]	Intergenic	AGATGCTGAAAGTGAGGCCATGTGGCTTGGCCAAATAAAACCTGGCTCCGTGGTGCCTCT[G/A]TCTTAGCAGCCACCCTGCTGATGAACTGCCACCTTGGACTTGGGACCAGAAAGAGGTGGG
84	rs8192288	[T/G]	Intergenic	ACACACTCAGGCTGCAGAGTCCCAGCCTTGTCACTTCCTATCTGTGTACGCCTGAAGCAG[C/A]TGCACCACTCCCTGCACCTCCTTTTCATCATCTGTAAAATGGGGGTGTGAGGCTGTCCAC
85	rs470411	[G/C]	Intergenic	GGAAGGTGATCGCTCCCCAAGCTCCCCAGGCCCTCACTCAGGACCTGACTACAGGCAACC[C/G]CGTCTGCTGAGGTCCTGCCCCTCTCCAGCCTCAAACTCAGTTGCACAGGTTCCATAAAGG
86	rs9394047	[A/C]	Intergenic	AGGGAGAGAGGGCAGGAGACCAAAGCAGAAGAGGAGCCCTAGAAGGAGGGCAGGAGCTGA[A/C]TGGGTCTGAAAATTTGTCTCAGAAAGCACAGGGACTCCCGTGTGCAGGGGCTGCCCTGGG
87	rs11809289	[T/C]	Intergenic	CGAAACACTTTTATCCACCCCCATTTATCTCTTCCTGACTCATGATCATCCTACAAAACT[A/G]TATTGGGCTTCACCTCCTCCCAGATTTCCTTCTCTCCTGCTCCTGTCCGGATTGTCGGCT
88	rs12885261	[T/C]	Intergenic	TCACCTATTTACAACAAAGGCACTATCTCAATTAGGGGGGAAGGGTAGTCTTTTCAACAC[G/A]TGCTGCTGGGAAACAAAATGACCCTTAACCTCTACAAACAAAAATTAGTTCAAAATGGAT
89	rs2297518	[A/G]	Synonymous	GTAGGACAACGGAAAAAGCTTTACCTGAATTTGTTGTTGAGCTCTTTCAGCATGAAGAGC[A/G]ATTTCTTCAGTTTCTAGAAAGAGAGGGAATGACAGAGTTCTCAAGCCAGGATGAATAAAA
90	rs1695	[A/G]	Synonymous	CAGGAGGCAGCCCTGGTGGACATGGTGAATGACGGCGTGGAGGACCTCCGCTGCAAATAC[A/G]TCTCCCTCATCTACACCAACTATGTGAGCATCTGCACCAGGGTTGGGCACTGGGGGCTGA
91	rs2305480	[T/C]	Synonymous	CACGCGCTTCTACCAAGACCCCAGCAGCATTAAAAAGGCTGCTTAGGAGAGGCTTGTCTG[A/G]GTCCTCCATGTGTAGCTCCCCGGAAATCAGGACCTCAGATACCTAGGGCCAGGAAGTGTG
92	rs2781659	[A/G]	Intergenic	CATGAGTTCACCAAGAACTGGACCTGTCAAGGTCAGCCCATCAACTTTGACAGGGAAGTC[A/G]GATTGGCAGAAGAAAAGGTAGTAAAGTGTGATGAGATTTTATTGAGACACCGGATAAATT
93	rs366631	[A/G]	Intergenic	TCGTTTTTATCAATTGGATCTAGACATCTTATCTTTCACAGCTCAAGACCCATTAACTCA[G/A]AATCATAAACTCTTAATGCATAATGAGAAATATAATGATTCCTAGGGCCAGGCACTTGTG
94	rs6599689	[T/C]	Intergenic	TCCCCTGTACAAAGCCAGCAAGGAAATAATTGAAGCTTTGTGGGCCCTGCAGTCTCTAAA[A/G]GCATGCTCACCTCTGCTGCCACAGCACAAAAGCAGCTGTTGAGAAAATGTCAGTGAGTGG
95	rs1051740	[T/C]	Synonymous	CATCCCTCTCTGGCTGGCGTTTTGCAAACATACCTTCAATCTTAGTCTTGAAGTGAGGGT[G/A]TCTGTTGAGAATCTCCACCTGCTTCTTCCAGTCAAATTCATTCCGCCAGTAGGAGATGAC
96	rs6017870	[T/C]	Intron	AAATGGGCAAAATACTTGAACAGGCACGTCACACACACCCAATATCAAAATGGACAATAG[G/A]TATAGAAAAGGTGCTCAGCCTCATCAGTAATTATGAGAAATCGAATCCCAAAATGACATA

Appendix E

Control Dashboards for All Samples in the Study

Illustration of controls dashboards, which is part of quality control procedure is presented in **Figure E1** for control dashboard of allele-specific extension, while **Figure E2** provides information on PCR Uniformity, the efficiency test of extension gap is shown in **Figure E3**. **Figure E4** demonstrates sample sex verification, while **Figure E5** and **Figure E6** represent first and second hybridization controls.

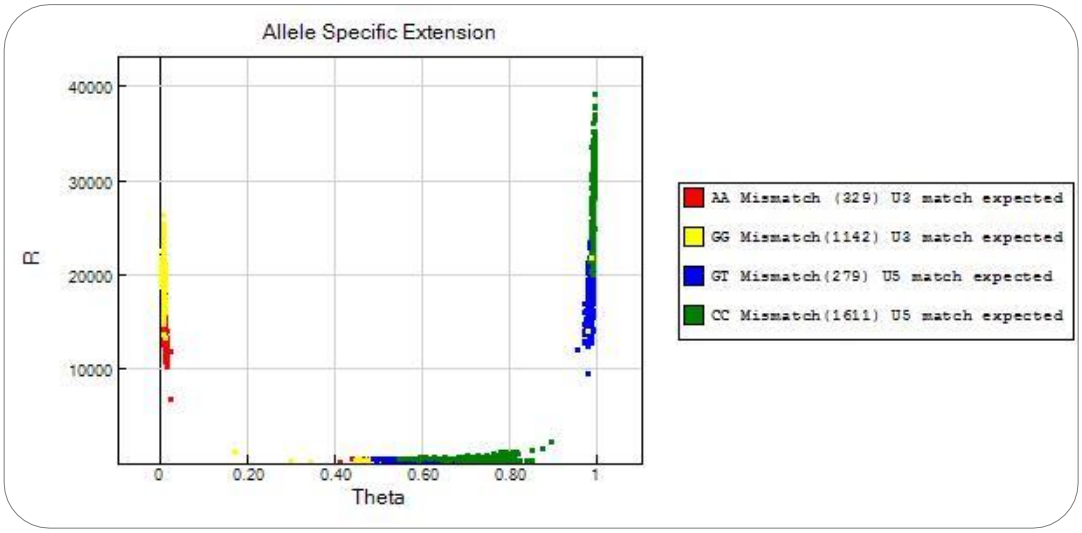


Figure E1 Allele Specific Extension (sample dependent). This control tests how effectively a correctly bound oligo is extended. AA (red) and GG (yellow) mismatches both have U3 and are detected with Cy3, whereas GT (blue) and CC (green) mismatches both have U5 and are detected with Cy5. Cy3 has a Theta value of 0, whereas Cy5 has a Theta value of 1 here. The important feature of this control is where the points lie on Theta, and the two controls that are detected together generally overlap. Based on the samples table in the GenomeStudio, the allele-specific controls for plates 5 and 11 have no intensities and thus failed. This pattern was observed in all controls. This control failure was of concern, but while looking at the good call rates, it was obvious, that this was a manufacturing issue with the controls (possibly not added to the OPA tube). If the call rates were very low, then failed controls would be an indicator of another underlying issue. (Personal Communication with Ingrid Du Plooy, Illumina Technical Support).

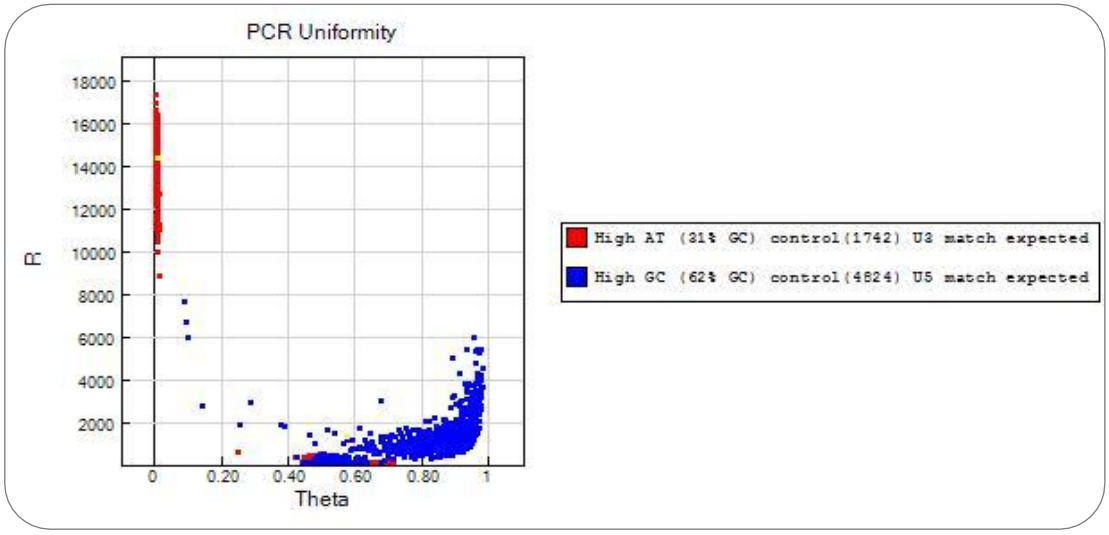


Figure E2 PCR Uniformity (sample dependent). The control tests the efficiency of the PCR amplification for high AT and GC regions. I found no GC bias in my samples, as the high GC controls show a lower signal. We would normally expect the GC signal to be slightly below the AT signal. It is a bit lower than normal; however, it is not expected to cause any problem.

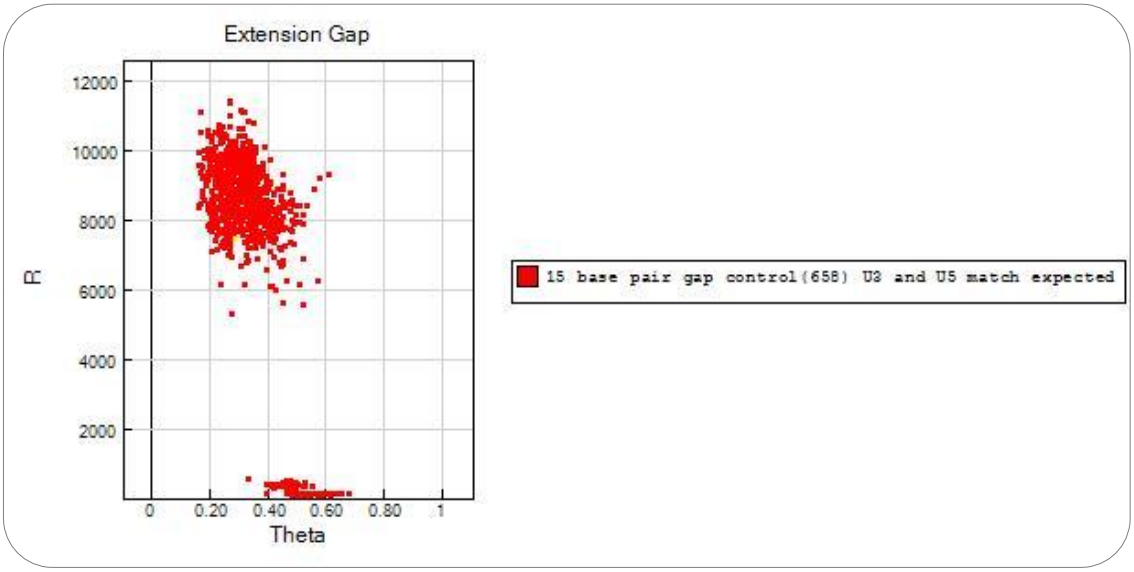


Figure E3 Extension Gap (sample dependent). The extension gap control measures the efficiency of filling the gap by 15 bases extension between the 3'-end of the allele specific oligo (ASO) and the 5'-end of the locus specific oligo (LSO). The failed controls were those from plates 5 and 11 and can be explained as outlined above in the Allele Specific Extension section. There are some controls with slightly lower intensities, but these are for samples with both high and slightly lower call rates, and there is no true separation from the main cluster.

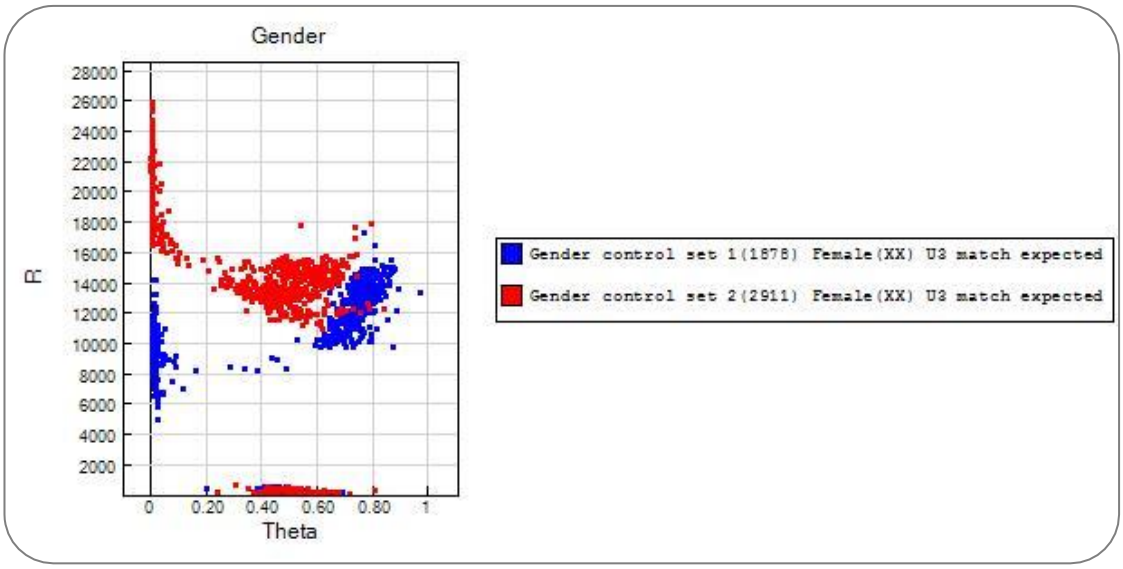


Figure E4 Gender (sample dependent). The gender estimate function in GenomeStudio works off the assay data from the gender controls. Essentially, there are 2 bead types (probes) on the array related to gender, therefore, each sample should be represented twice for each probe. The bead intensities for the gender control assays are used to compute a theta value for these loci. There is no specific cut off for the theta values at which clusters for male and female samples should appear in the Gender Control panel. However, the gender controls in GoldenGate are designed against human sequences such that females (XX) should be in the homozygous cluster at small theta values (usually <0.2), and males (XY) should be in the heterozygous cluster around 0.6. The theta value is based on intensity in green vs red channel. Overall, the Gender control dashboard looked fine; the Homozygote (XX) samples are ~0 and Heterozygote samples clustering at ~0.6. However, 40 samples demonstrated mismatch gender based on the reported sex and inferred sex. Again, the samples with low R (intensity) were from plate 5 and 11 and therefore gender could not be determined for these samples.

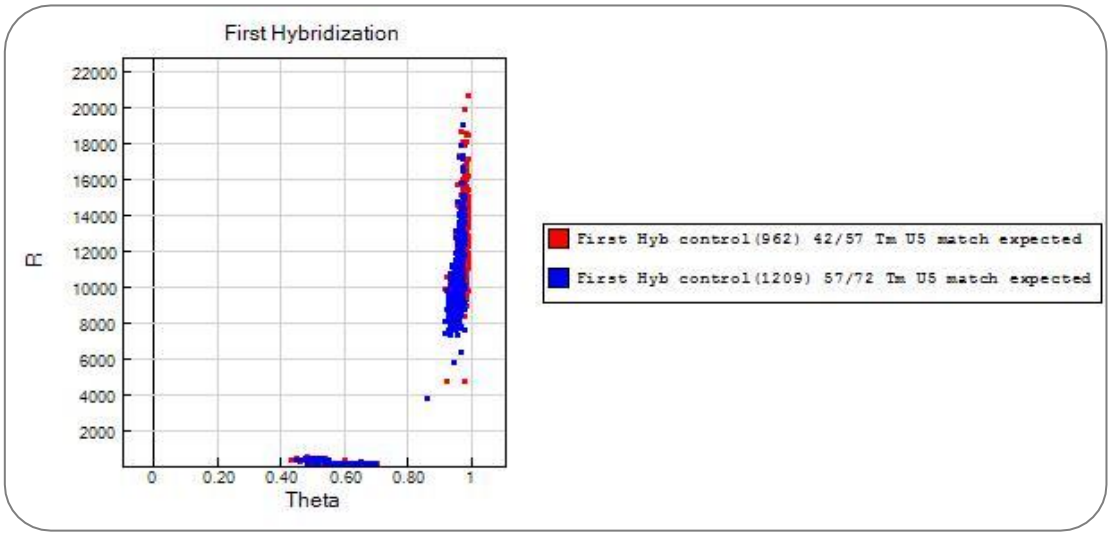


Figure E5 First Hybridization Controls (sample dependent). This control investigates the effect of various melting temperatures (Tm) on the efficiency of annealing allele specific oligos (ASOs) to the specific DNA locus. The compatibility between the two signals would decide a Cy5 match and therefore we expect the signal to be Theta ~1. Overall, both the oligos bound (Cy5 match) and exhibit good signal with single clustering. If they had failed we might have expected one or both to give no or a low signal, or for there to be a significant difference between the two signals. Samples with low intensity reflect poor quality of DNA. Once again, the zero intensity represents plates with absence of control oligos. (Personal Communication with Marianne Mitchell, Illumina Technical Support).

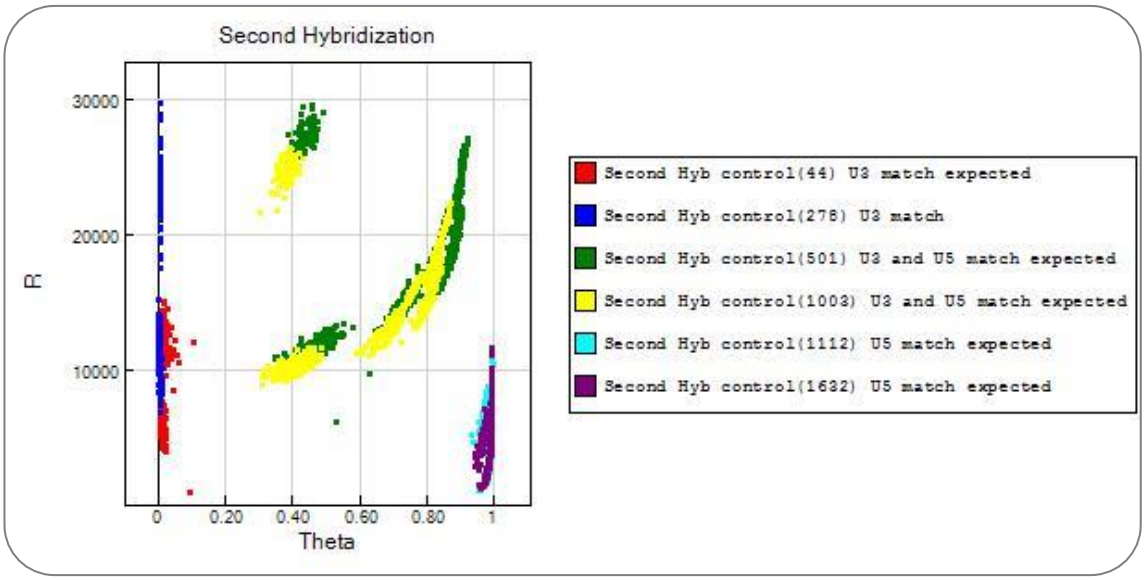


Figure E6 Second Hybridization Controls. This is a sample independent test, which signifies the capability of the single stranded PCR products to hybridize to the array beads. These oligos are designed to give signals as if they were homozygotes (both AA and BB) and heterozygotes (AB). Therefore, the important feature here is that the red and blue, which showed fluoresce in the green channel, have a Theta value of ~0. Whereas, the yellow and green, which showed fluoresce in both channels, are around Theta ~0.5 (it is fine if this is shifted a bit as long as it is clearly distinct from the homozygous groups) and the blue and the purple which showed fluoresce in the red channel have a Theta ~1. Based on the plot, it appears that data originate from three different batches, thus, we may find several clusters at different positions, each representing the data from one batch. These probably different scanners were used for scanning, had different parameters settings and maybe there were slight differences in the way the plates were processed. Nevertheless, this does not affect the hybridization process. Since, samples that are dropping out in the sample dependent controls seem not affected with regard to call rates, this does not seem to be a sample quality issue, but rather related to the absence of controls oligos (Personal communication with Susanne Angelow, Illumina technical support).

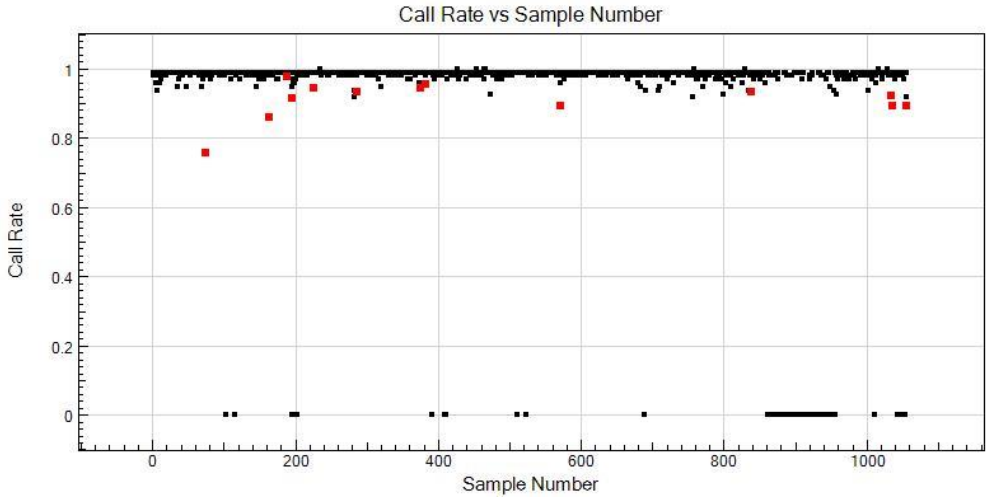
Appendix F

Samples Evaluation for Genotyping Quality Control

Samples were assessed to identify any poorly performing samples to improve the quality of the genotype dataset in general. The examination was performed based on the call rate and GenCall (GC) for each sample. **Figure F1** illustrates raw data of all samples before exclusion of any SNPs. Each sample is represented by one data point. Viewing the call rate of all samples quickly reveals poorly performing samples with abnormally low call rates (red highlighted). Subsequent to exclusion of samples with call rates <0.9, the dataset improved greatly as in **Figure F2**.

Figure F3 shows raw data of all samples based on GenCall before reprocessing of samples. The poorly performing samples are in the lower left quadrant (blue highlighted samples were re-amplified) **Figure F4** demonstrates raw data of 988 samples before exclusion of any samples and SNPs. The poorly performing samples are in the lower left quadrant (red highlighted samples were removed from final dataset).

1)



2)

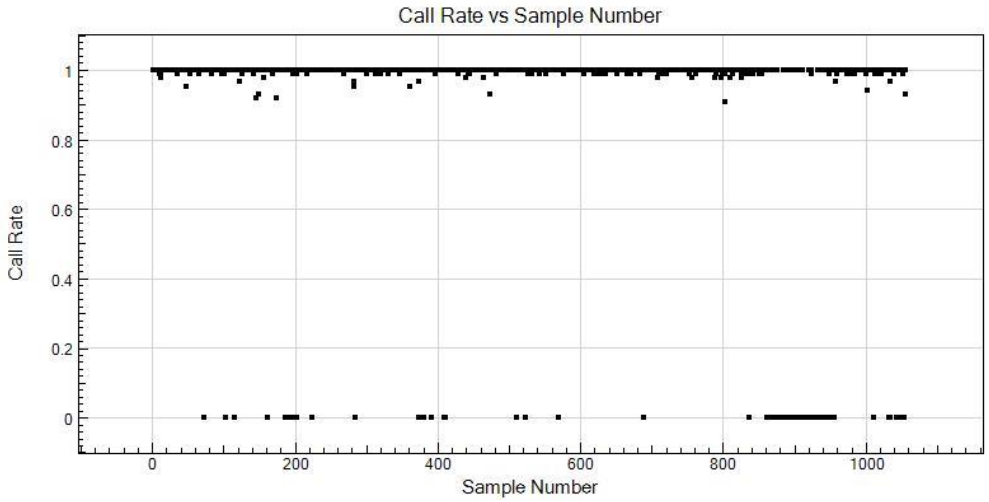
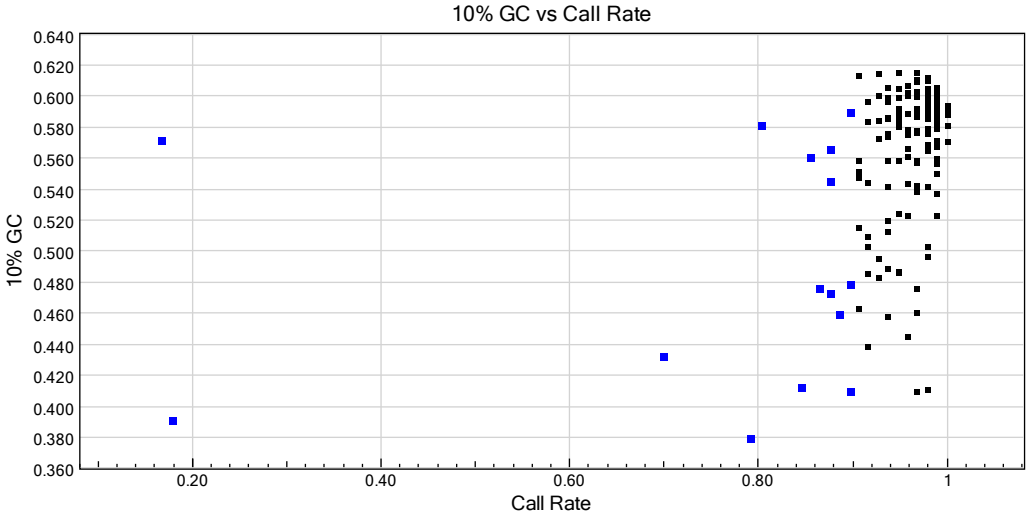


Figure F1 Scatter plot of call rates across the set of samples before exclusion of any SNPs.

Figure F2 Scatter plot of call rates across the set of samples subsequent to exclusion poor performing samples.

3)



4)

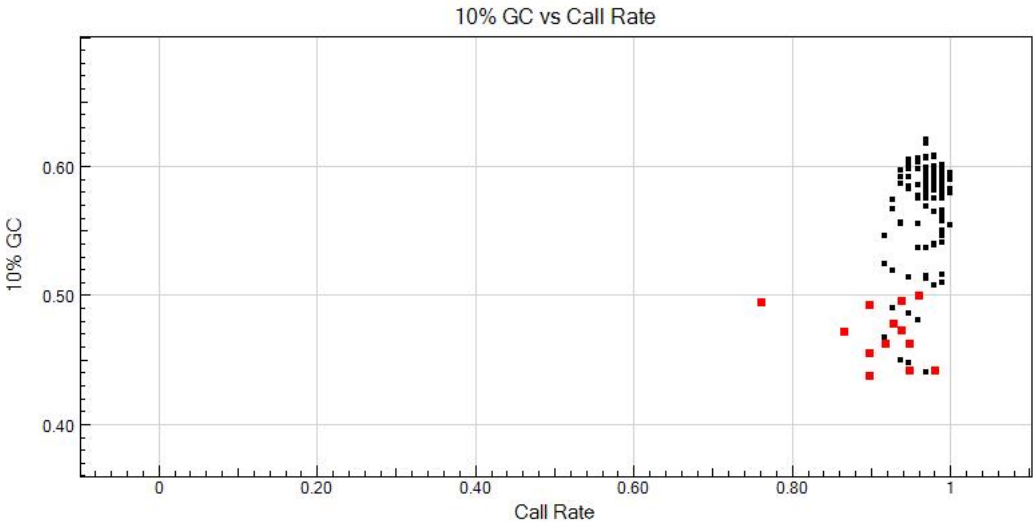


Figure F3 Scatter plot of 10% GC score compared to call rates of the sample set before reprocessing of samples. Figure F3 Scatter plot of 10% GC score compared to call rates of the sample set before exclusion of any samples and SNPs. Each sample is represented by one data point. (Note: Y-axes does not start at the origin).

Appendix G

SNPs Evaluation

Figure G1 demonstrates comparison of original GenTrain score prior and subsequent to manual editing for each SNP. Overall, GenTrain score has improved following the manual editing, which reflected good cluster separation. In general, all the SNPs have GenTrain score above 0.60. The zero score implies SNPs, which have failed (zeroed), therefore, have been excluded from the final dataset.

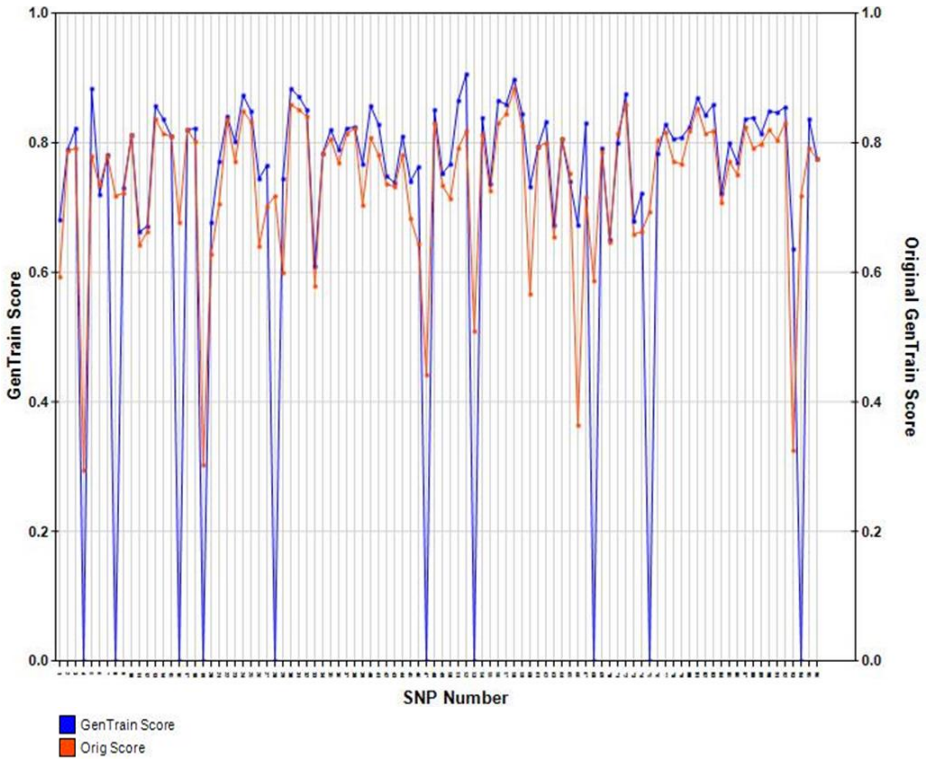


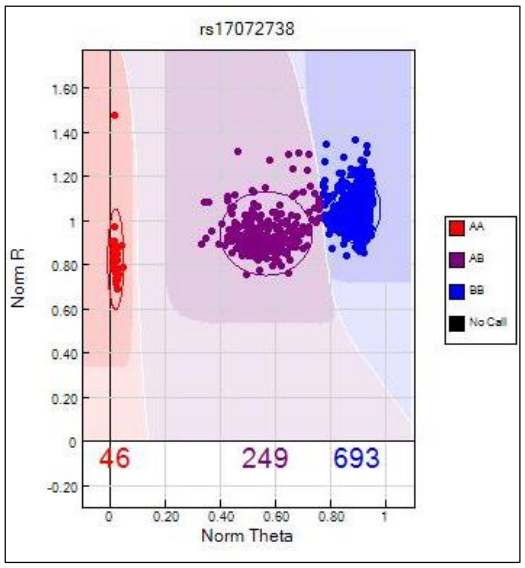
Figure G1 Line plot of edited GenTrain score and original GenTrain score vs SNP number. Whereas **Table G1** provides details of cluster separation, SNP frequency, GenTrain score and other distribution scores for all 96 SNPs. SNP graph for each failed SNPs was based on the overlap cluster (**Figure G2**), low SNP frequency (**Figure G3**) and low AB R mean (**Figure G4**). Some of the successfully genotyped SNPs prior to editing and subsequent to manual editing are shown in **Figure G5** and **Figure G6**.

Table G1 SNP Table with Quality Check Distribution

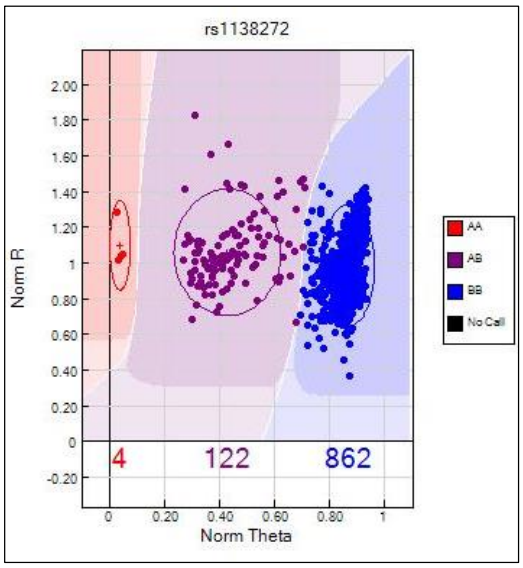
SNP Index	SNP ID	No Calls	Calls	Call Frequency	Gentrain Score	Cluster Separation	AB T Mean	AB R Mean
1	rs2917666	0	974	1	0.6813	0.3349	0.62	1.828
2	rs7782389	0	974	1	0.7903	0.3201	0.727	0.556
3	rs997279	1	973	0.999	0.8235	0.5387	0.451	1.222
5	rs3759757	1	973	0.999	0.8839	0.4863	0.469	1.353
6	rs1137933	1	973	0.999	0.7218	0.2535	0.187	1.099
7	rs2826003	0	974	1	0.7816	0.4403	0.666	1.542
9	rs2246012	1	973	0.999	0.7313	0.3447	0.76	0.886
10	rs2074113	4	970	0.996	0.8129	0.3705	0.793	0.904
11	rs7232792	1	973	0.999	0.6635	0.2968	0.854	1.319
12	rs2284659	2	972	0.998	0.6717	0.4614	0.604	1.408
13	rs2781668	0	974	1	0.8571	0.4809	0.41	1.211
14	rs2066853	1	973	0.999	0.8384	0.4775	0.704	1.229
15	rs2733262	4	970	0.996	0.8103	0.3432	0.707	0.803
17	rs17599586	1	973	0.999	0.821	0.3745	0.711	1.264
18	rs742869	3	971	0.997	0.8233	0.3572	0.315	1.092
20	rs2364723	0	974	1	0.6775	0.4999	0.549	1.322
21	rs8078340	4	970	0.996	0.7713	0.3502	0.226	1.141
22	rs3742879	0	974	1	0.8423	0.399	0.642	1.1
23	rs268691	2	972	0.998	0.8034	0.2916	0.256	1.318
24	rs2301022	3	971	0.997	0.8739	0.529	0.562	1.229
25	rs3170633	1	973	0.999	0.8499	0.4357	0.616	0.533
26	rs2779251	2	972	0.998	0.7452	0.2471	0.193	1.328
27	rs7132743	2	972	0.998	0.7666	0.2673	0.755	1.021
29	rs6770096	0	974	1	0.7446	0.3468	0.848	1.263
30	rs7557529	3	971	0.997	0.884	0.4946	0.563	0.908
31	rs2779253	1	973	0.999	0.8723	0.4846	0.619	1.875
32	rs2606345	0	974	1	0.8518	0.403	0.346	1.215
33	rs8192287	4	970	0.996	0.6103	0.5094	0.616	1.042
34	rs2781666	0	974	1	0.7841	0.3131	0.691	1.085
35	rs4795080	6	968	0.994	0.8202	0.3997	0.629	0.549
36	rs7144243	4	970	0.996	0.7898	0.3599	0.743	1.071
37	rs17722841	1	973	0.999	0.8221	0.4165	0.742	1.092
38	rs699473	2	972	0.998	0.826	0.339	0.313	0.504
39	rs1889022	1	973	0.999	0.7671	0.2527	0.312	0.651
40	rs9282799	0	974	1	0.8585	0.5415	0.534	1.472
41	rs7591449	10	964	0.99	0.8283	0.3787	0.288	1.343
42	rs3894194	1	973	0.999	0.7487	0.4008	0.356	0.805
43	rs1442293	1	973	0.999	0.7383	0.2919	0.205	0.449
44	rs749174	0	974	1	0.8107	0.4045	0.595	1.079
45	rs689453	0	974	1	0.7417	0.2609	0.18	0.853
46	rs3796644	2	972	0.998	0.7631	0.2482	0.78	0.679
48	rs17779352	1	973	0.999	0.8521	0.5223	0.32	1.11
49	rs2128238	6	968	0.994	0.753	0.2624	0.629	1.084

SNP Index	SNP ID	No Calls	Calls	Call Frequency	Gentrain Score	Cluster Separation	AB T Mean	AB R Mean
50	rs1799814	10	964	0.99	0.7688	0.3999	0.589	0.53
51	rs10454231	9	965	0.991	0.8666	0.518	0.474	1.096
52	rs7776785	1	973	0.999	0.9069	0.7308	0.454	1.608
54	rs7140310	1	973	0.999	0.8398	0.4932	0.483	1.11
55	rs1800566	0	974	1	0.7376	0.4178	0.75	1.074
56	rs17861115	0	974	1	0.8665	0.4979	0.6	0.835
57	rs11735827	0	974	1	0.859	0.4379	0.425	0.923
58	rs2531866	0	974	1	0.8978	0.5444	0.471	0.732
59	rs9900426	2	972	0.998	0.8465	0.4513	0.573	1.101
60	rs10517	1	973	0.999	0.7337	0.2662	0.827	0.725
61	rs2282885	4	970	0.996	0.7946	0.4164	0.727	0.539
62	rs2001350	1	973	0.999	0.8324	0.602	0.622	0.834
63	rs527705	2	972	0.998	0.6737	0.2436	0.828	0.892
64	rs4902503	1	973	0.999	0.8067	0.4395	0.311	0.426
65	rs7875663	1	973	0.999	0.7408	0.3105	0.828	1.115
66	rs407257	2	972	0.998	0.674	0.4977	0.543	1.567
67	rs2531872	3	971	0.997	0.8316	0.3528	0.321	1.379
69	rs221454	2	972	0.998	0.7918	0.3568	0.718	1.162
70	rs2781667	2	972	0.998	0.6509	0.4547	0.347	1.668
71	rs13306703	3	971	0.997	0.8012	0.4032	0.289	1.374
72	rs2198843	1	973	0.999	0.8767	0.4932	0.418	1.196
73	rs778233	1	973	0.999	0.6798	0.4963	0.527	1.643
74	rs7156352	6	968	0.994	0.7238	0.506	0.416	1.354
76	rs3824781	2	972	0.998	0.7849	0.4882	0.346	0.536
77	rs2236687	3	971	0.997	0.8299	0.4657	0.672	0.949
78	rs10459953	0	974	1	0.8073	0.4723	0.603	1.191
79	rs10483801	7	967	0.993	0.808	0.5269	0.496	0.221
80	rs2153747	2	972	0.998	0.8253	0.3516	0.277	1.301
81	rs10853181	0	974	1	0.87	0.5048	0.423	1.049
82	rs1014025	3	971	0.997	0.8433	0.4709	0.582	1.091
83	rs2779248	0	974	1	0.8598	0.4616	0.414	1.118
84	rs8192288	0	974	1	0.7223	0.3726	0.534	0.27
85	rs470411	0	974	1	0.8009	0.3529	0.195	1.135
86	rs9394047	1	973	0.999	0.7701	0.4106	0.36	1.229
87	rs11809289	1	973	0.999	0.8369	0.4776	0.295	1.226
88	rs12885261	1	973	0.999	0.8394	0.4813	0.619	1.054
89	rs2297518	1	973	0.999	0.8148	0.4615	0.387	0.428
90	rs1695	1	973	0.999	0.8499	0.4442	0.376	0.758
91	rs2305480	1	973	0.999	0.8479	0.5163	0.529	0.963
92	rs2781659	3	971	0.997	0.8567	0.5253	0.407	1.286
93	rs366631	1	973	0.999	0.6365	0.2877	0.254	1.407
95	rs1051740	7	967	0.993	0.8368	0.4643	0.599	1.073
96	rs6017870	1	973	0.999	0.7765	0.3029	0.2	1.052

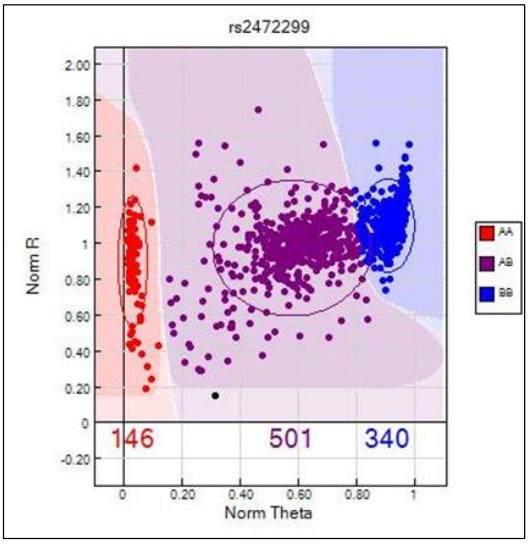
Random SNP (SNP 8)



GSTP1 Exon 6 Ala114Val (SNP 16)



CYP1A1 T/C (SNP 68)



EC-SOD3' UTR T/C (SNP 53)

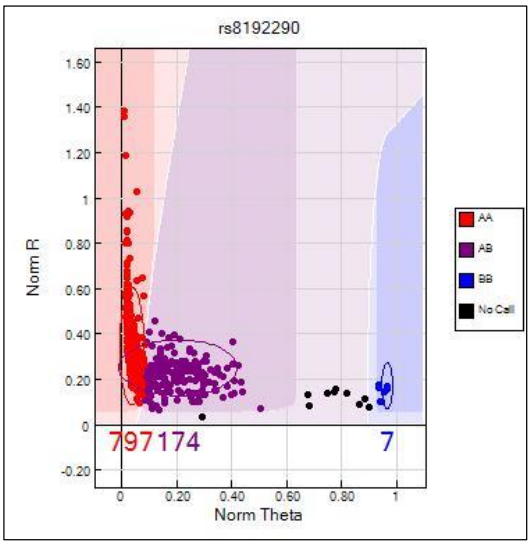
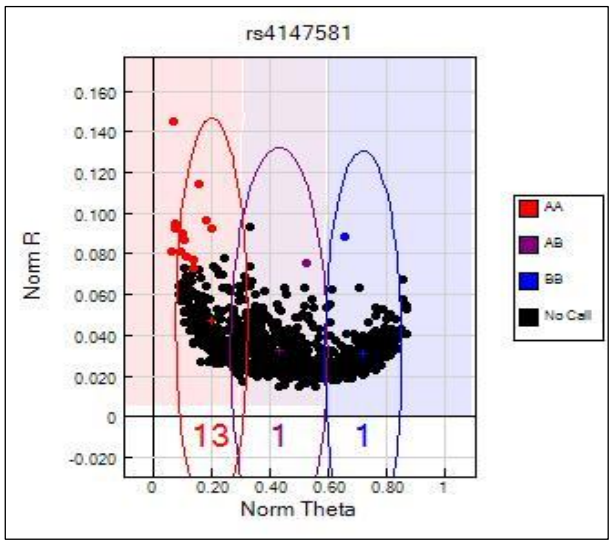


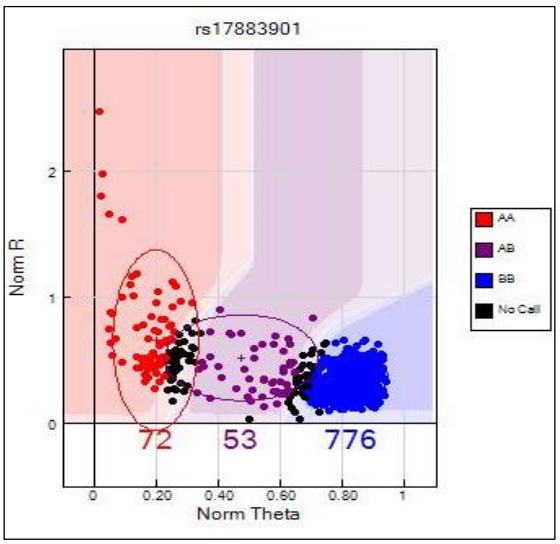
Figure G2 SNPs with Overlap Clusters. The AB, BB and AA clusters undergo overlapping, which makes the genotyping unreliable.

GSTP1 C/G (SNP 4)



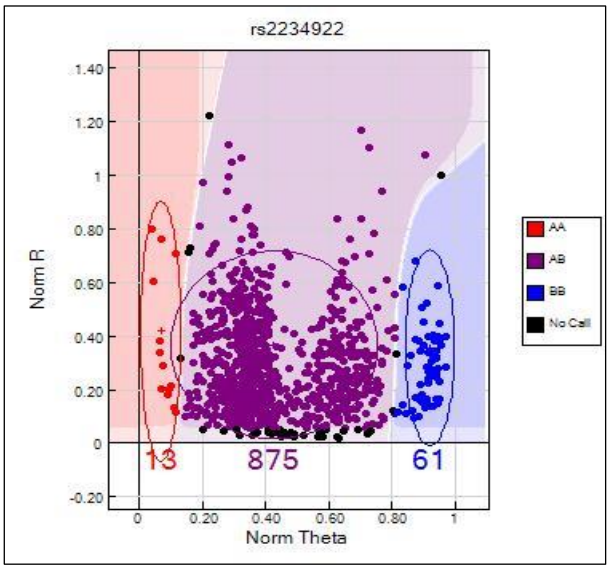
This SNP has too low intensities, overlap clusters, and low SNP frequency.

GCLC A/G (SNP 19)



Diffused clusters. Significantly high number of unresolved "no call" samples and poorly separated clusters produces unreliable genotyping data.

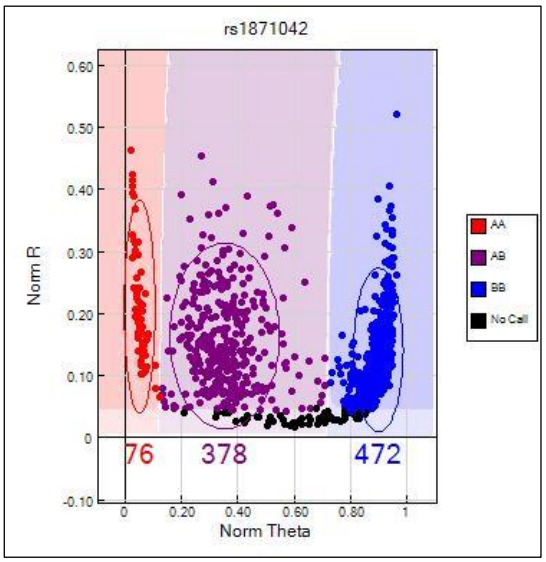
EPHX1 His139Arg (SNP 47)



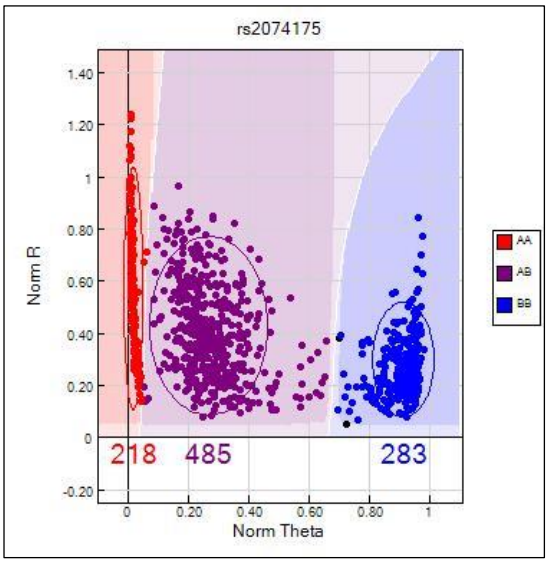
This SNP has excess of heterozygosity, therefore, this causes problems in manual editing of the locus for correct genotyping call.

Figure G3 SNPs with Low SNP frequency and High Heterozygosity Ratio.

GSTP1 C/T (SNP 28)



Random SNP G/A (SNP 75)



Random SNP A/G (SNP 94)

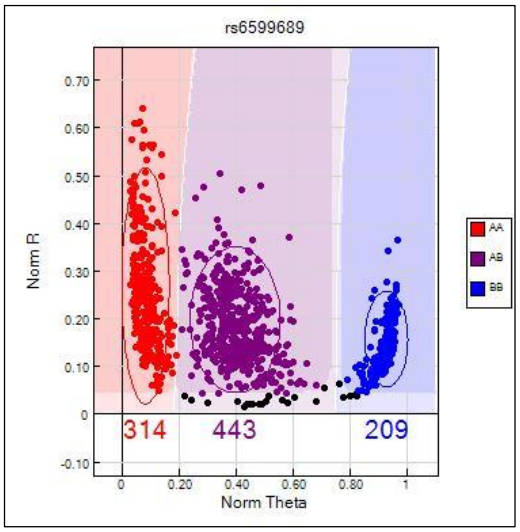
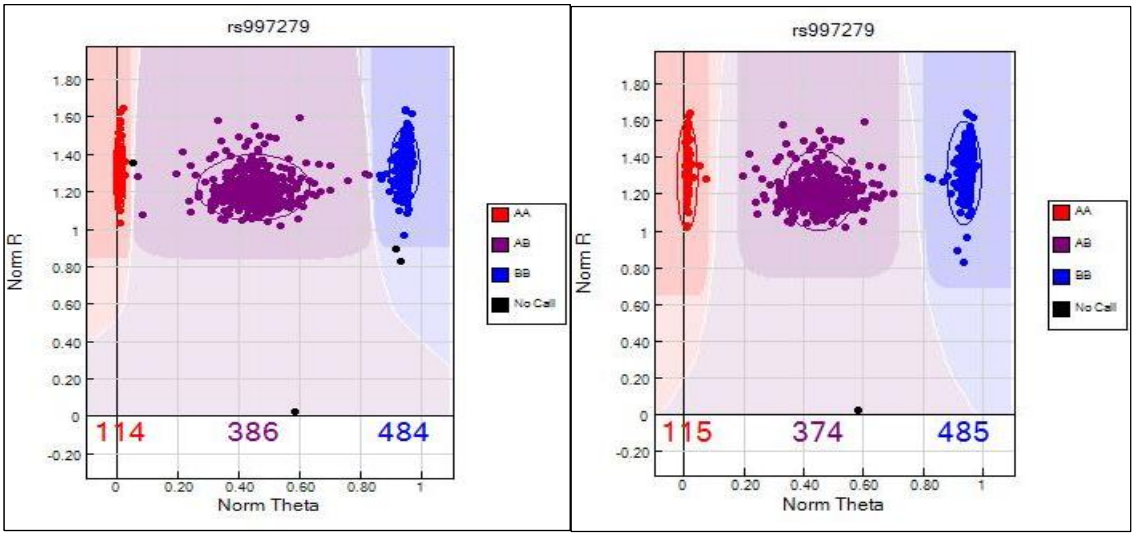


Figure G4 SNPs with Low AB R Mean. Very low intensities (Norm R). The normalized R values of clusters are extremely low for the successful manual editing of the genotypes leading to unreliable calling.

Before editing **After editing**

SNP 3



SNP 5

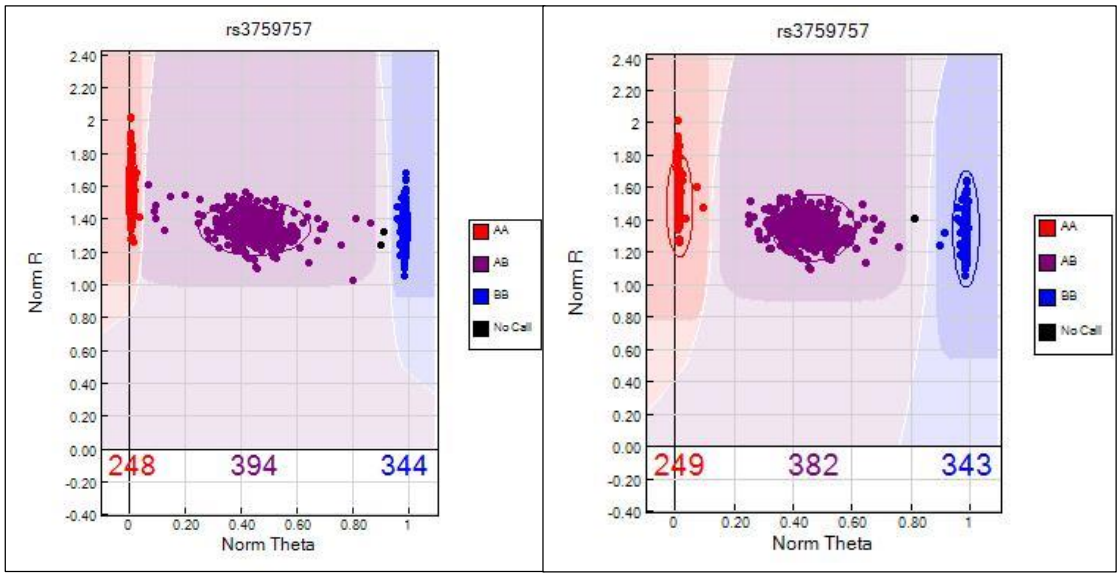


Figure G5 SNP graphs for successfully manually edited SNP3 and SNP5. In the right panel (top and bottom) are the SNPs, after the editing. Exclusions of outlier samples for SNP3 and SNP5 and correcting genotype call improve call frequency for both SNPs.

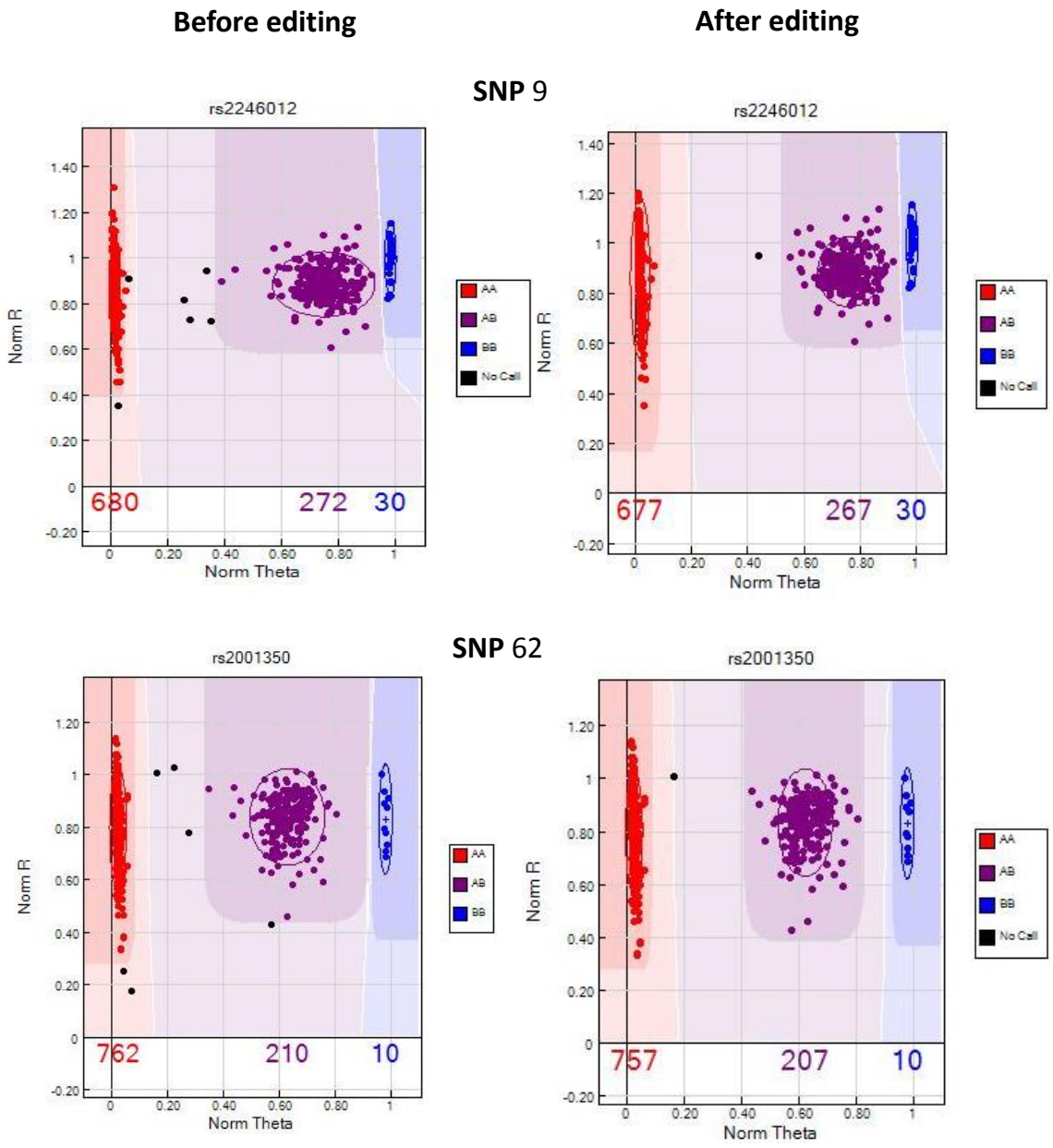


Figure G6 SNP graphs for successfully manually edited SNP9 and SNP62. Following exclusion of poorly performing samples, SNPs cluster were manually edited. In the right panel (top and bottom) are the SNPs, after the editing. More samples were included for AA and AB clusters to improve call frequency of the SNPs. Three distinct clusters of AA, AB and BB were observed.

Appendix H

Population Structure Analysis

In the current study, a unique set of SNPS were selected without prior knowledge of their function in stratification of population to act as potential ancestral informative markers (AIMS). One marker was identified on the long and short arm of each autosome using computer random number generator to identify the genomic position. Only SNPs not associated with specific diseases and showing the greatest frequency variation across the HapMap populations (within 20kb of the selected position) were included in the analytical SNP set – further detail of the selection criteria are provided below. Our approach was to determine whether we could obtain sub-division of ancestral populations, based on a limited set of genetic loci. This represents an alternative to employing established AIMS, which have been used to stratify between population continental and inter-continental origins, where 100's to 1000's of genetic loci are generally employed (Kidd et al., 2011; Kosoy et al., 2008; Tian et al., 2008).

The AIMS selection criteria included the following: (1) **Informativeness (In)** - This measure was applied to identify highly informative markers with large allele frequency differences between populations. Thus, the AIMS set were evaluated using analysis of Fst to determine the ability of the markers to distinguish between various sub-population group. (2) **Linkage disequilibrium** - The SNP selection also requied a uniform genome-wide distribution and to minimal linkage disequilibrium. The markers with highest In, with a minimum inter-SNP distance >500 Kb, were selected and any SNP with strong LD ($r^2 > 0.8$) was removed (Tian et al. 2008). These aspects are described in detail below:

A measure of the divergence of each subpopulation from their ancestral origin is usually measured by Fixation indices statistics (Fst) (Wright 1965). This describes the degree of differentiation in the variations of two alleles (heterozygosity) between the subpopulations. It also explains that the degree of the ancestors of the two investigated alleles is similar and samples are taken from unrelated subpopulations. The Fst measurement is usually described as a single value based on the variance of allele frequencies on a locus deviated from the overall allele frequency of the subpopulations using the following formula:

$$F_{ST} = \frac{\mathrm{Var}(p)}{\bar{p}(1 - \bar{p})}$$

F_{ST} equals to the variance of the allele frequencies of the subpopulations divided by the expected mean allele frequency (heterozygotes) across all subpopulations (Holsinger and Lewis 2002).

It is possible to identify the level of deviation in the evolution of subpopulations only when the average Fst involving a vast number of multiallelic loci is taken into consideration. Therefore, in STRUCTURE the Fst that refers to the F model was applied to multiallelic loci based on correlated allele frequencies model. This model was adopted from Nicholson et al. (2002) and assumes that individuals in a population diverged from the same "ancestral population" at different rate and that therefore the ancestral allele frequencies are correlated to the subpopulations. The allele frequencies of closely linked subpopulations, are likely to be related and therefore form the basis of the correlated frequencies model (Falush et al. 2003). Thus, the capability of STRUCTURE to identify related but diverged populations is significantly enhanced when the correlation model is added in the population structure model (Rosenberg et al. 2005). By making a minor modification in the traditional model, representation of the allele j at locus l in the correlated population by pAlj, represents p where the Fk characterizes Fst in the classical model. The use of a distinct F value such as F1, F2, F3, and FK for each population instead of one value for a range of populations indicates that

allele frequencies for each K population in the study population, have independently deviated from ancestral allele frequencies at specific rates (Falush et al. 2003).

$$\mathrm{Var}(p_{klj}) = F_k p_{Alj}(1 - p_{Alj})$$

Generally, small Fst values (close to zero) represent similar allele frequencies within the distinct population whereas large Fst values (close to 1) indicating large variation in the allele frequencies. Thereby, we would thus expect that an increase in Fst signifies the lack of migration in natural population, which is reflected by the higher subdivision amongst the main population and the populations become static for distinct alleles (Holsinger and Weir 2009). Conversely, significantly admixed populations offer challenge in accurate determination of the Fst values. The Fst values would be greater at a locus that shows preference over alleles under the natural selection, compared to the Fst values at a locus where genetic variation amongst population is mainly due to genetic drift (Guo et al. 2009).

Table H2 shows the STRUCTURE simulation summary for Fst, likelihood probability and variance values for all the chains. Most of the cluster populations produced similar likelihood at different run at each K values. On the other hand, similarities observed in the probability values in all of the 5 chains denote that convergence was achieved by undertaking sufficient iterations. In addition, some of the distinct and most occurring patterns observed during the single run analysis were rather consistent (see **Figure H1**). A plateau for the log probability values in the STRUCTURE attained between the K values of 3 to 5, suggesting that most possible of the clusters fall in this category of K values. The highest value of LnP(D) (-28261.3) was obtained for K=3 of pattern B (Run 8) which indicates the best fit for the sampled populations and it is the most common pattern seen.

From these analyses there appeared to be a marked segregation between the black and the Eurasian (European/White & Asian) populations at all of the K values, suggesting little or insignificant admixture between the two populations, with the first

recognizable major split observed at K=2. However, the North African (Population 4: Sudanese, Moroccan, Somali) possesses population structure intermediate between Asian and the African populations. This is observed in all patterns at each K value. The population 'Other' in the ethnicity classification, consisting of mixed and difficult to define ethnicities showed a gradient mirroring the Eurasian cluster. Additionally, there was some uniformity in the clustering of the populations across the K=2 to K=5.

Overall, there was no clear separation between the clusters in Eurasian populations at K=3 (see **Figure H2**). On the contrary, the black population (blue) presented significantly visible genetic structure at K=3, distinguishing them from the other populations and again suggesting little admixture. The European (Population 6) population demonstrated uniformity of cluster formation, with a gradient across European population to Asian populations, ending with the East Asia & South East Asian (Population 8). The rest of Asian population consisted of South Central Asian (Population 7: Afghan, Indian, Nepali, Pakistani and Bangladeshi), and other Asian background (Population 9). Noticeably, the South East Asian (Chinese, Filipino, Hong Kong Chinese and Vietnamese) shows distinct red cluster compared to all of the Asian subpopulations. The South West Asian population (Population 5: Kurdish, Turkish and White Turkish) did not form separate cluster but was broadly in line with the European whites. When the population structure was analysed at higher K values, K=4 and K=5, the cluster formation was more complex with some enhanced division between the Eurasian and 'Other' populations (see **Figure H1**).

Table H1 Individuals sampled in the study and sample size for each of the 11 subpopulations

Subpopulations Code	Subpopulations Detail	Main Ethnicity	Self-reported Ethnicity	No. of Individuals
1	Africa	Black	African, Caribbean	150
2	West Africa	Black	Ghanaian, Guinea, Nigeria, Sierra Leon, Togolese	66
3	South & Central Africa	Black	Angolan, Congolese	12
4	North & East Africa	Black	Sudanese, Moroccan, Somali	19
5	South West Asia	Others	Kurdish, Turkish, White Turkish	36
6	Europe	White	English, Albanian, Portuguese, Traveller	267
7	South Central Asia	Asian	Afghan, Indian, Nepali, Pakistani, Bangladeshi	316
8	East Asia & South East Asia	Asian	Vietnamese, Filipino, Chinese, Hong Kong Chinese, Oriental	26
9	Other Asia	Asian	Any other Asian background	5
10	America	Others	Brazillian Latin	11
11	Other Mixed	Others	Any other mixed, Mixed White and Asian, Mixed White and Black, Mixed Black and Asian, Other	66
			Total Number of Children	**974**

Individuals are grouped in subpopulations, according to self-reported ethnicity. Subpopulations are ordered according to geographical region beginning from Africa, to South West Asia (Middle East), then Europe to Asia, followed by America and ends with "Others Mixed" populations. The entire mixed ethnicity groups were grouped in one subpopulation, "Others", as there are very few of them in each subgroup. Asia is the major subpopulation in the study, which consists of ~35% of the investigated populations, followed by 27% of Europe subpopulation, and 25% of Africa subpopulations. The 'main ethnicity' category is used for the ethnicity classification in the association analysis whereas the 11 subpopulations category was used in the Structure analysis.

Table H2 shows the simulation analysis representing likelihood, variance, Alpha and Fst values. Structure was run 10 times for each K value (K=2 to K=5). The run with highest likelihood (-28261.3) is highlighted in red and was obtained for K=3, which represent the best fit for the sampled populations.

Table H2 STRUCTURE Simulation Summary

K	Run Name	LnP(D)	Var[LnP(D)]	α1	Fst_1	Fst_2	Fst_3	Fst_4	Fst_5
2	Run 1	-28277	433.5	0.1095	0.16	0.1203	-	-	-
2	Run 2	-28271.1	429	0.1201	0.1248	0.1588	-	-	-
2	Run 3	-28294.5	471.6	0.1162	0.1148	0.1663	-	-	-
2	Run 4	-28297.8	466.6	0.102	0.1537	0.1203	-	-	-
2	Run 5	-28282.4	444	0.1089	0.1364	0.1406	-	-	-
2	Run 6	-28285.9	453.2	0.1104	0.1624	0.1201	-	-	-
2	Run 7	-28310.9	487	0.0989	0.2083	0.0486	-	-	-
2	Run 8	-28285.8	441.7	0.1004	0.1302	0.1405	-	-	-
2	Run 9	-28280.1	434.3	0.1033	0.159	0.117	-	-	-
2	Run 10	-28287.4	463.1	0.1227	0.1283	0.1556	-	-	-
3	Run 11	-28263.1	804.4	0.0796	0.0016	0.0535	0.263	-	-
3	Run 12	-28269	813.5	0.0779	0.0536	0.0007	0.2627	-	-
3	Run 13	-28270.1	816.3	0.0813	0.0523	0.0042	0.2631	-	-
3	Run 14	-28274.2	826.3	0.0825	0.265	0.0535	0.0016	-	-
3	Run 15	-28261.3	798.2	0.0773	0.0046	0.2603	0.0522	-	-
3	Run 16	-28290.5	856.3	0.0799	0.263	0.002	0.053	-	-
3	Run 17	-28276.3	829.4	0.0797	0.0532	0.0023	0.2641	-	-
3	Run 18	-28275.4	822.9	0.0773	0.0024	0.2605	0.0524	-	-
3	Run 19	-28270	815.8	0.0789	0.0516	0.0046	0.2622	-	-
3	Run 20	-28279.9	830.8	0.0769	0.0519	0.2608	0.0038	-	-
4	Run 21	-28905.4	2167.6	0.0664	0.0168	0.2745	0.0329	0.0272	-
4	Run 22	-28723.8	1771.5	0.0646	0.041	0.2731	0.0035	0.0253	-
4	Run 23	-28784.1	1872.8	0.0657	0.2765	0.0318	0.0144	0.0186	-
4	Run 24	-28789.9	1941.5	0.0655	0.0273	0.0241	0.2734	0.0245	-
4	Run 25	-28886.7	2139	0.0678	0.0165	0.0207	0.041	0.2757	-
4	Run 26	-28913.2	2177.2	0.0628	0.0273	0.2718	0.0238	0.0254	-
4	Run 27	-28905.8	2168.5	0.0644	0.0319	0.0273	0.0175	0.2742	-
4	Run 28	-28726.2	1770.7	0.0644	0.0361	0.0042	0.2753	0.0268	-
4	Run 29	-28873.6	2087.9	0.0669	0.0079	0.0376	0.0264	0.2761	-
4	Run 30	-28904.8	2176.4	0.0699	0.0271	0.0374	0.0137	0.2767	-
5	Run 31	-28976.1	2349.5	0.0594	0.0361	0.2877	0.0158	0.0265	0.0149
5	Run 32	-29167.4	2724.6	0.0627	0.0159	0.0189	0.2905	0.0303	0.0295
5	Run 33	-28987.4	2374	0.0616	0.2892	0.0231	0.0141	0.0142	0.0434
5	Run 34	-29127.8	2644.7	0.0586	0.0401	0.0163	0.2878	0.0145	0.0197
5	Run 35	-28833.9	2019.8	0.0623	0.0108	0.0054	0.2924	0.0161	0.0504
5	Run 36	-28655.6	1651.4	0.059	0.0511	0.0087	0.2892	0	0.0207
5	Run 37	-29189.1	2736.8	0.0609	0.2873	0.0175	0.0028	0.039	0.0303
5	Run 38	-28906.5	2198.1	0.062	0.0114	0.0263	0.0135	0.2914	0.04
5	Run 39	-29172	2704.4	0.0587	0.0238	0.02	0.0001	0.0412	0.2888
5	Run 40	-28873	2111.6	0.0608	0.2895	0.0161	0.0203	0.0182	0.0325

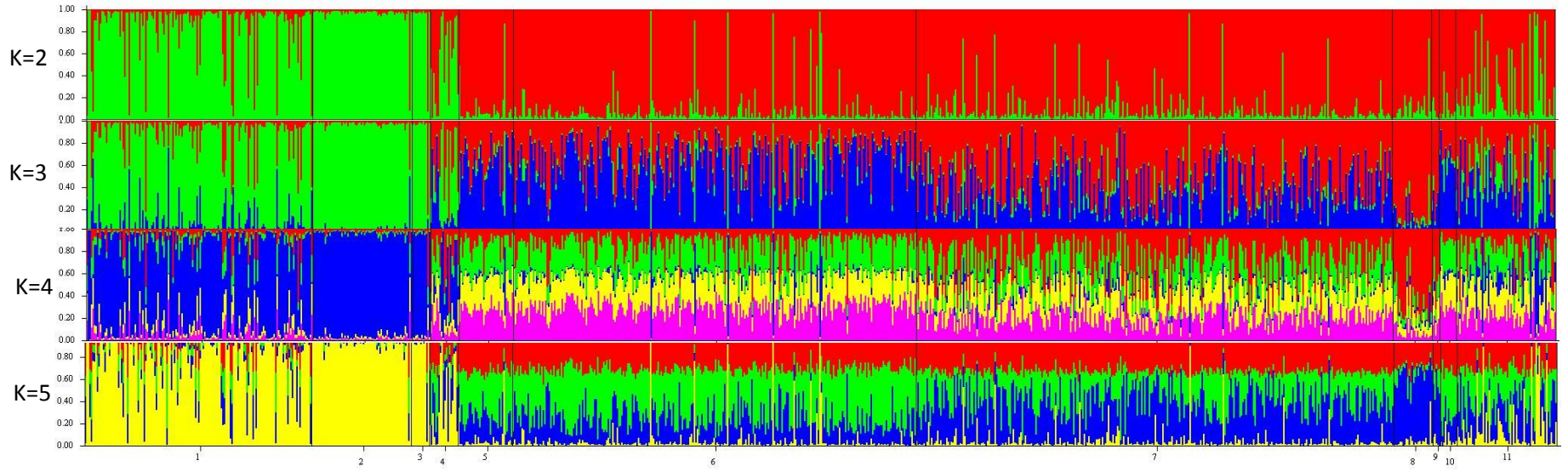


Figure H1 Structure Summary with Highest Likelihood in Each K Value (K=2 to K=5).

Illustration of structure pattern with the highest likelihoods for all K values with the run number out of ten runs. K represents number of different clusters simulated in the structure analyses. A distinct thin vertical line indicates each individual and these lines are subdivided into various coloured divisions of respective K clusters. Black lines represent segregation of various populations. Population labels are as in **Table H1**. A plateau for the log probability values in the STRUCTURE attained between the K values of 3 to 5, suggesting that most possible of the clusters fall in this category of K values. There appears to be a marked division between the black and the Eurasian populations at all of the K values, suggesting little or insignificant admixture between the two populations and the first recognizable major split was observed at K=2. When the population is analysed at higher K value, K=4 and K=5, the cluster formation is more complicated and rather subdivides the existing admixed pattern observed in Eurasian and 'Other' populations.

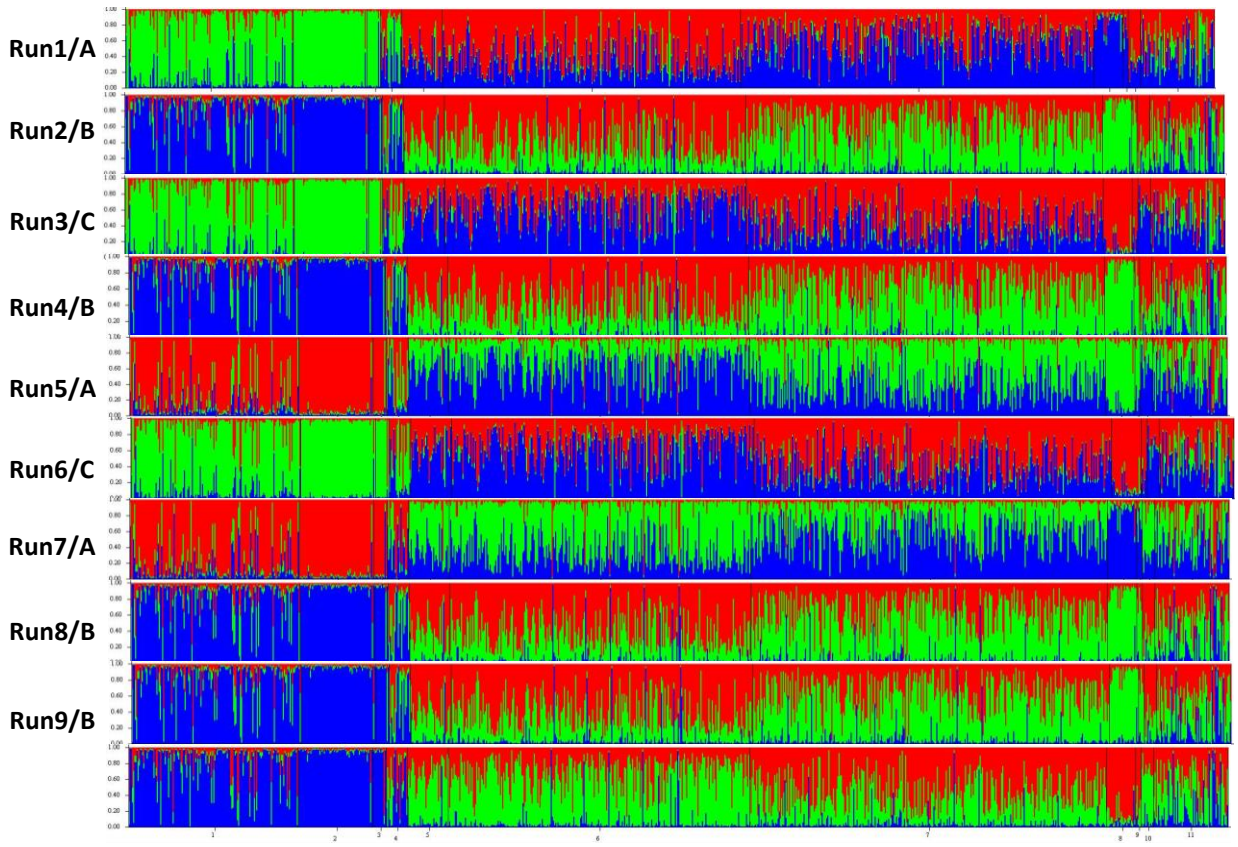


Figure H2 Population Structure Pattern at K = 3.

Illustration of structure pattern for the best possible K value in the study populations. Line graphs indicate likelihood of each run while dashed lines represent actual K value. Bar plots are showing different clustering pattern observed across 10 runs in each K, in this case K=3. Vertical bars signify individuals in populations while solid lines segregate individuals according to subpopulations. Distinct colour indicates classification of populations in specific cluster. The highest likelihood is obtained for K=3 of pattern B (Run 8) which indicates the best fit for the sampled populations and it is the most common pattern observed overall. There is some distinct cluster for Europeans (mostly red) and Asian (mostly green); however, it is depicted as admixed at different proportion.

The mean Fst values based on the K=3 for cluster 1, 2 and 3 were 0.0046, 0.2603 and 0.0522 respectively. Here, Cluster 1 represents Asian, Cluster 2 represents mostly Black and Cluster 3 characterizes European subpopulations. The low Fst values severely affect the capability of the STRUCTURE to identify differences in related subpopulations. This was apparent for Asian and White populations, which is reflected in the pattern observed in the bar plots. There are several possibilities for the observed conditions. Fst determines the extent of allele frequency deviation in a sample taking into consideration the probability of most likely deviation possible. An estimate of 0.46% probably indicates that there are very few common alleles among the studied populations. Therefore a rare allele might lower the Fst value when low allele frequency is observed in the subpopulation divided by the allele frequency of the total population. Another condition could be that when the Asian and White population have similar allele frequencies (see **Table H3** for allele frequency for each SNPs investigated in the STUCTURE analysis), this does not clearly separate populations from each other, but demonstrates a high level of admixture. Therefore, it indicates little divergence amongst the populations. The allele frequency divergence among the Cluster 1 vs Cluster2, Cluster 1 vs Cluster 3 and Cluster 2 vs Cluster 3 were 0.0834, 0.0069 and 0.0927 respectively. This indicates that there is weak separation between Asian and White ethnicity/Europe subpopulations reflecting evidence observed in Fst value. Overall, average distances estimation (average expected heterozygosity for all loci) between individuals in the same cluster for cluster 1, cluster 2 and cluster 3 is 0.3664, 0.3454 and 0.3389 respectively.

The probability of population admixture depends on the mean Dirichlet parameter, denoted by the alpha (α). The high value of α (>1) suggests significantly higher admixture amongst the populations and lower α values (close to zero) categorise individuals to the distinct populations (Falush et al. 2003). For instance, at K=3 the mean Dirichlet parameter calculated for the highest likelihood is 0.0773 and is representative of the three major populations (cluster 1, cluster 2 and cluster 3) recognised in the clustering approach.

Table H3 Allele Frequency for Random SNPs based on dbSNP

SNPs Index	SNPs	Chr	European (CEU)		South Asian(GIH)		East Asian (HCB)		West African (YRI)		East African (LWK)	
			Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
2	rs7782389	7	0.513	0.487	0.523	0.477	0.651	0.349	0.712	0.288	0.534	0.466
3	rs997279	6	0.611	0.389	0.699	0.301	0.895	0.105	0.633	0.367	0.606	0.394
7	rs2826003	21	0.898	0.102	0.852	0.148	0.570	0.430	0.327	0.673	0.311	0.689
8	rs17072738	18	0.017	0.983			0.222	0.778	0.475	0.525		
11	rs7232792	18	0.814	0.186	0.744	0.256	0.547	0.453	0.659	0.341	0.600	0.400
15	rs2733262	1	0.544	0.456	0.676	0.324	0.767	0.233	0.168	0.832	0.206	0.794
23	rs268691	19	0.509	0.491	0.472	0.528	0.733	0.267	0.097	0.903	0.191	0.809
27	rs7132743	12	0.822	0.178			0.944	0.056	0.237	0.763		
29	rs6770096	3	0.898	0.102	0.812	0.188	0.756	0.244	0.699	0.301	0.800	0.200
41	rs7591449	2	0.118	0.882	0.364	0.636	0.605	0.395	0.580	0.420	0.694	0.306
43	rs1442293	15	0.759	0.241			0.544	0.456	0.892	0.108		
46	rs3796644	4	0.371	0.629	0.335	0.665	0.488	0.512	0.168	0.832	0.211	0.789
49	rs2128238	16	0.643	0.357			0.686	0.314	0.679	0.321	0.811	0.189
51	rs10454231	3	0.071	0.929	0.216	0.784	0.244	0.756	0.686	0.314	0.539	0.461
52	rs7776785	7	0.042	0.958			0.100	0.900	0.175	0.825		
57	rs11735827	4	0.951	0.049	0.892	0.108	0.977	0.023	0.819	0.181	0.800	0.200
59	rs9900426	17	0.442	0.558			0.356	0.644	0.733	0.267		
63	rs527705	9	0.544	0.456	0.426	0.574	0.628	0.372	0.805	0.195	0.811	0.189
65	rs7875663	9	0.558	0.442	0.540	0.460	0.581	0.419	0.314	0.686	0.228	0.772
69	rs221454	14	0.420	0.580	0.477	0.523	0.337	0.663	0.956	0.044	0.978	0.022
73	rs778233	5	0.535	0.465	0.670	0.330	0.429	0.571	0.274	0.726	0.264	0.736
75	rs2074175	19	0.469	0.531	0.602	0.398			0.243	0.757		
76	rs3824781	10	0.095	0.905	0.034	0.966	0.155	0.845	0.054	0.946	0.103	0.897
77	rs2236687	21	0.500	0.500			0.573	0.427	0.283	0.717		
80	rs2153747	13	0.863	0.137	0.920	0.080	0.667	0.333	0.518	0.482	0.661	0.339
85	rs470411	11	0.271	0.729			0.678	0.322	0.325	0.675		
86	rs9394047	6	0.942	0.058	0.801	0.199	0.814	0.186	0.929	0.071	0.961	0.039
87	rs11809289	1	0.992	0.008			0.878	0.122	0.658	0.342		
94	rs6599689	10	0.482	0.518	0.648	0.352			0.296	0.704		
96	rs6017870	20	0.788	0.212			0.875	0.125	0.392	0.608		

13 SNPs	SNPs which don't clearly separate between CEU (European) and GIH (South Asian) or East Asian when the GIH populations was not genotyped
2 SNPs	SNPs which don't clearly separate between main ancestry (CEU/GIH (or HCB when the GIH was not genotyped)/YRI)
9 SNPs	SNPs which have different MAF between HCB (East Asian) and GIH (South Asian)
2 SNPs	SNPs which don't clearly separate CEU (European) and YRI (African)
3 SNPs	Failed SNPs (excluded from the final dataset)

CEU= Utah residents with Northern and Western European ancestry
GIH = Gujarati Indians in Houston, Texas
HCB= Han Chinese in Bejing, China
YRI = Yoruba in Ibadan, Nigeria
LWK= Luhya in Webuye, Kenya

Triangle plot of all 11 subpopulations were shown in **Figure H3** whereas **Figure H4** shows the triangle plot at K=3 representing the clustering behaviour involving individuals of the three major populations (Asians=red coded), (Whites=blue coded) and (Black=green coded) as observed in the bar plot. Unsurprisingly, the Asians and White failed to form independent clusters and instead clustered at the edge of the major cluster, which is formed of rest of the individuals (admixed). On the other hand, majority of the black population (green) appeared to be clustering at the vertex of the triangle plot although a small percentage appeared to be scattered throughout the triangle area particularly at the centre of the triangle plot. Individuals with missing genotype data, which are coded with red colour, can be seen at the centre of the plot. The marked difference observed in some populations signifies possibilities of population stratification; however, investigation of more loci warrants differentiation of the Asian population from the White individuals.

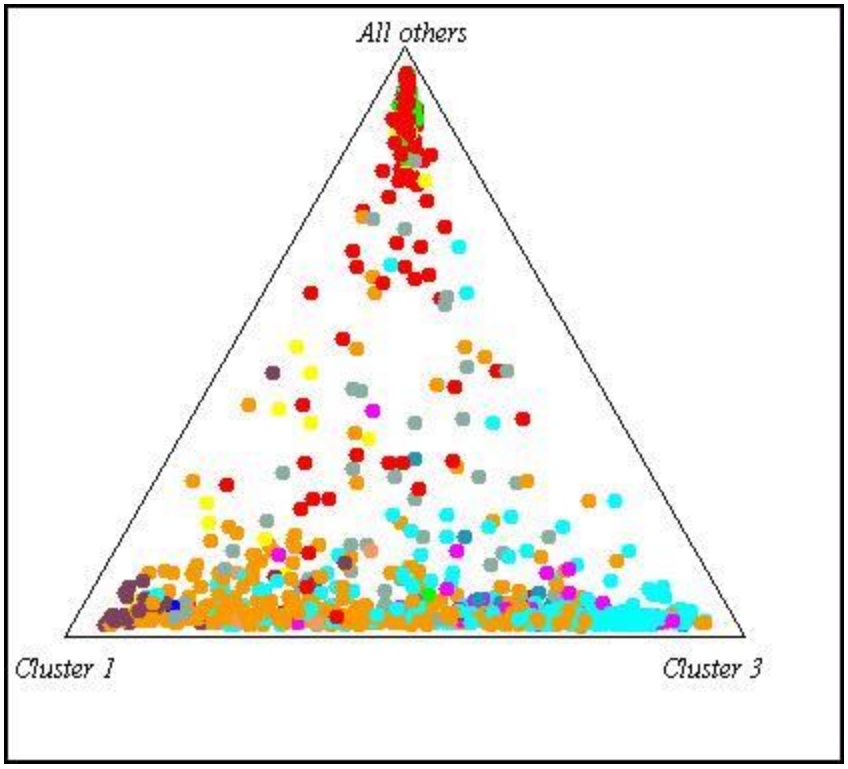


Figure H3 Triangle Plot at K = 3. Triangle plot represents each 11 subpopulations analysed in the population structure. Each subpopulation coded with different colour for identification. Cluster 1, cluster 2 (All Others) and cluster 3 denote Asian, Black and White ethnicity in the population study.

- (Pop 1) Africa
- (Pop 2) West Africa
- (Pop 3) South & Central Africa
- (Pop 4) North & East Africa
- (Pop 5) South West Asia
- (Pop 6) Europe
- (Pop 7) South Central Asia
- (Pop 8) East Asia & South East Asia
- (Pop 9) Other Asia
- (Pop 10) America
- (Pop 11) Other Mixed

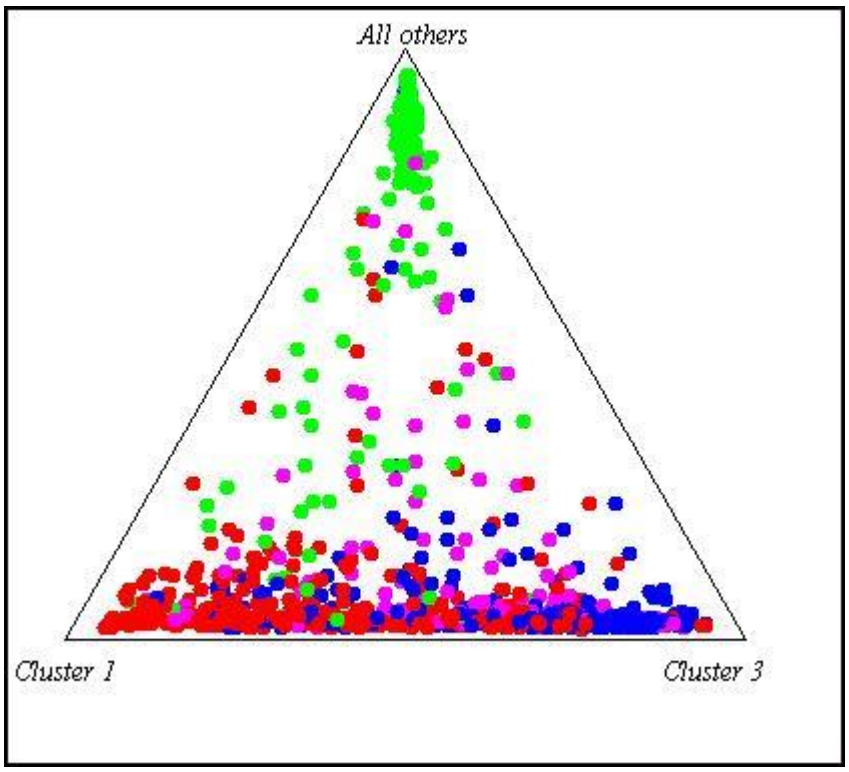


Figure H4 Triangle plot showing results of STRUCTURE analysis for K=3. Each point represents an individual for the specific cluster based on the colour code. There is no true separation for Asian and White and both are rather shown as admixed populations (Eurasian).

Cluster 1: Asian
Cluster 2: Black
Cluster 3: White
Other Mixed ethnicity

It is important to have strong population structure because with weak population structure they appear admixed at K=2 causing individuals to fall in to unspecified clusters with characteristics similar to all of the tested clusters as observed in the present data. The degree of admixture amongst the populations depends on the population sample size. Differences in the sample size of the individual populations also affect the allele frequencies of the said cluster and the differentiation (if any) would be representative of the

population with large number of individuals compared to the population with the small number of individuals. This will lead to the scenario where small population flanked by large number of cline populations appearing to be admixed at the most extreme areas of the cline (Kidd et al. 2011). This is reflected at K=3 for North & East Africa population (Sudanese, Moroccan & Somali) and South West Asian (Turkish) populations. Our data set consist of ~3%, ~4% and ~2% of East Asian & South East Asian, South West Asian, and North & East African populations respectively of the total sample. Whereas the major subpopulation, which is South Central Asian, consists of 32.4% of the whole population studied. In addition, at higher K values, the small subpopulation flanked by bigger subpopulation that appeared admixed, have the potential to form a single cluster due to estimation of the extreme clusters of the cline. This is illustrated in Population 2: West African and Population 8: East Asian & South East Asian at K=5. It should be noted that none of the subpopulations appear as clear cluster except for African subpopulations (exceptional for Population 4). Therefore, determination of K value for the populations investigated was rather an ad hoc approximation, thus, allele frequencies for a particular cluster to be suggestive of their ancestors may not essentially be true. In addition, allocation of subpopulations to more than one cluster could not be inferred as true admixture. Moreover, it is highly likely that deviation of data from HWE will result in K value being unreliable. Although, deviation from the HWE would significantly affect the closely related population but it hardly has any effect on the unrelated populations.

There appears to be a lack of discrete structure in the current dataset. However, an increase in the subjects in each population or increased markers could potentially enhance chances of identifying any discernible structure in the data. As observed in the data, the Eurasian and African populations showed poor cluster separation. This suggests that the random markers in our panel do not stratify the Central, South and East Asian origin effectively. Probably, the inclusion of large number of SNPs with significantly high level of allele frequency variations relative to Eurasians could result in greater segregation of the west and east Eurasian populations. This set of SNPs also does not optimally segregate

even the populations amongst the individuals of the East Asian origin. A similar outcome was shown in Kidd et al. (2011) study. However, we should bear in mind that the SNPs employed in this analysis were selected randomly without prior knowledge of the function in structure analysis. These SNPs have never been tested in any other studies. Therefore, we could propose the SNPs panel selected in this study to best distinguish between White and African. As South West African (Somali) cline with Asian population, therefore, it is recommended not to employ this panel on this group.

Global population structure is complicated by intermediate populations such as the South West Asian, South East Asian and the West African. At high frequency of these populations, it leads to the structure and genetic differences cline instead of clustering. In addition, major population groups such as Africans, Caucasians and Asians are rather heterogeneous and further divided into complicated subpopulations. The statistical power of the study significantly depends on the sample size and any change in the population size would affect the overall deviation. An increase in the sample size may cause small allele frequencies variation in the samples compared to the allele frequencies in the primary population (Holsinger and Weir 2009). On the other hand, it is impossible to control the deviation in the populations related to the underlying genetics, which is an inherent genetic material in evolutionary process leading to diversity between populations. This leads to the conclusion as before that K values heavily depend on the population sample size and individuals sampled in the study population.

Appendix I

Supplementary Data for Chapter 4

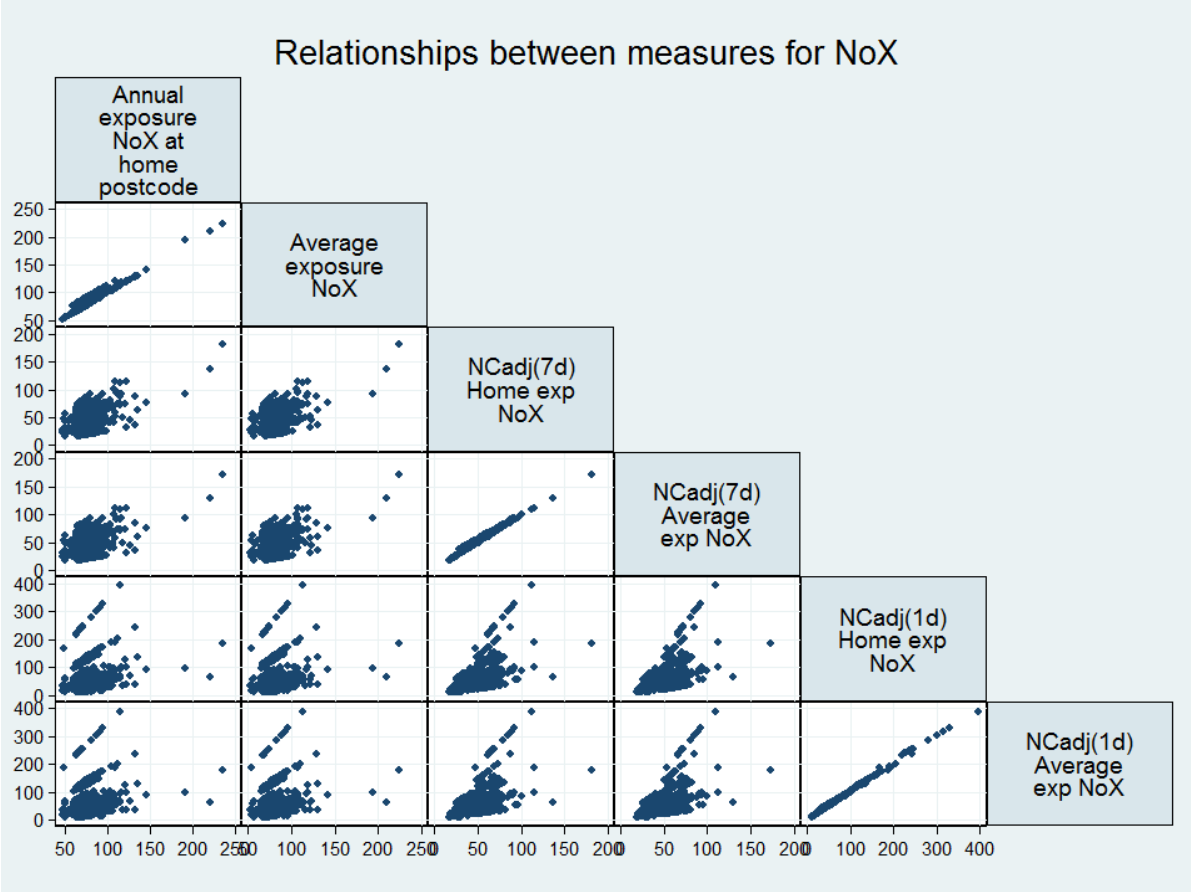


Figure I1: Correlations between chronic, sub-chronic and acute exposure NOx attributions modeled using 20m radius buffers around each subject residential address (home), or weighted for periods spent at the home and school location (average).

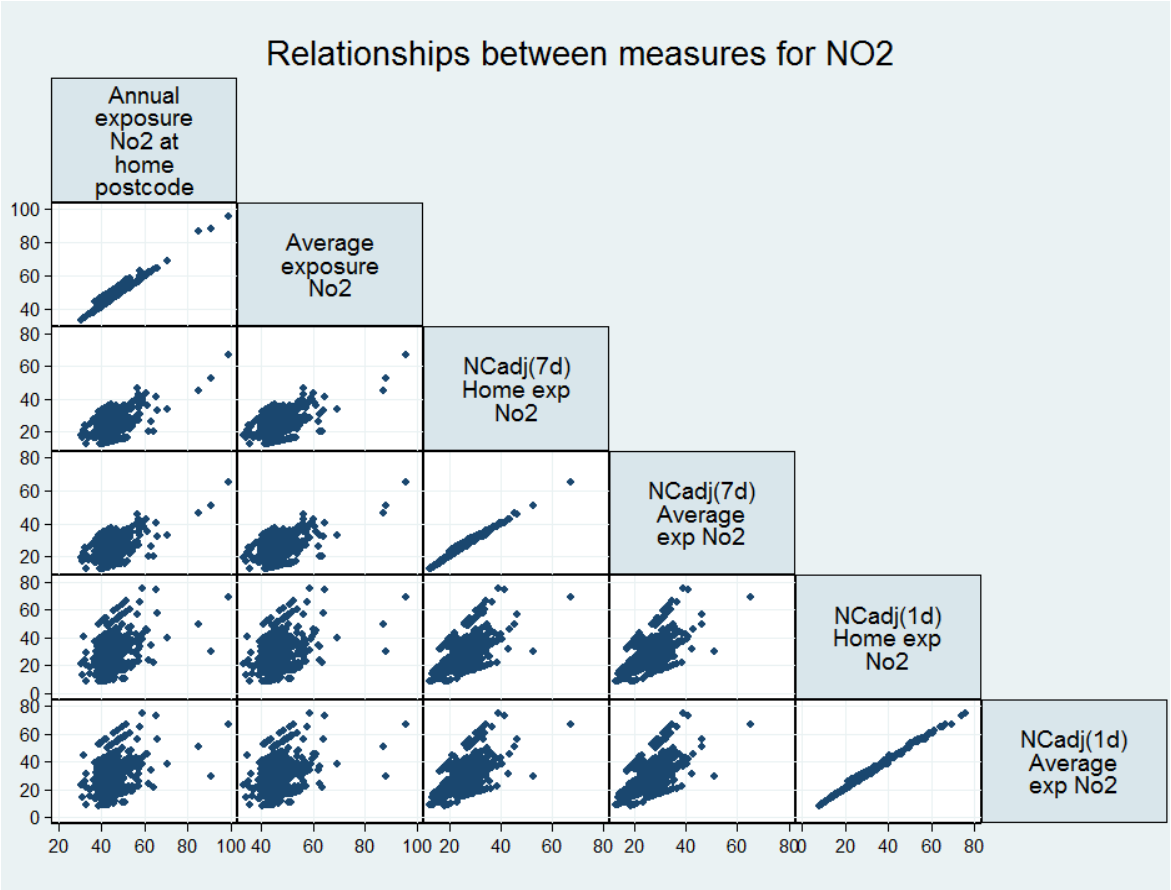


Figure I2: Correlations between chronic, sub-chronic and acute exposure NO_2 attributions modeled using 20m radius buffers around each subject residential address (home), or weighted for periods spent at the home and school location (average).

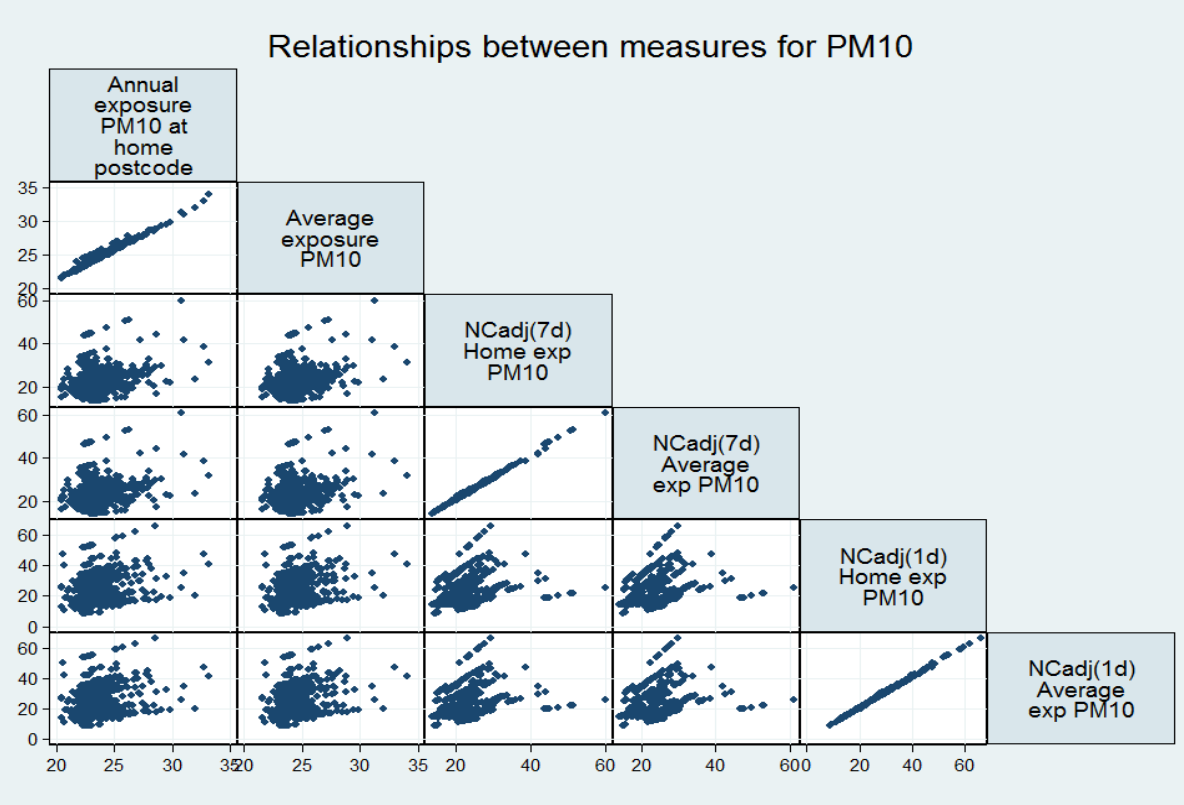


Figure I3: Correlations between chronic, sub-chronic and acute exposure PM_{10} attributions modeled using 20m radius buffers around each subject residential address (home), or weighted for periods spent at the home and school location (average).

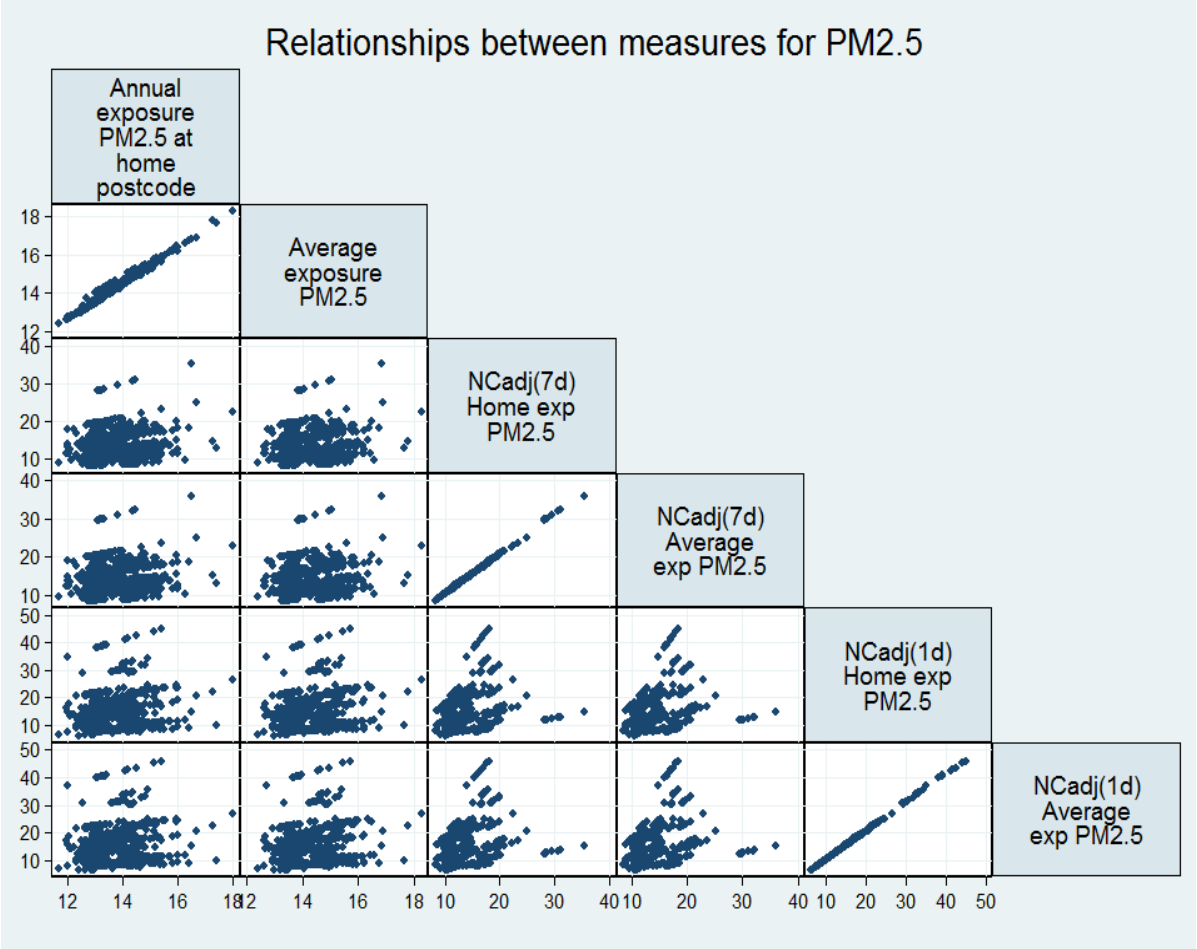


Figure I4: Correlations between chronic, sub-chronic and acute exposure $PM_{2.5}$ attributions modeled using 20m radius buffers around each subject residential address (home), or weighted for periods spent at the home and school location (average).

Table I1: Hardy-Weinberg equilibrium (HWE) p-values for SNPs in GSTM1, GSTP1 and NQO1 genes by ethnicity

SNP ID	Asian	Black	White	Others
GSTM1				
*rs366631	-	-	-	-
GSTP1				
rs749174	0.28	0.86	0.79	1.00
rs1695	0.89	0.12	0.59	1.00
NQO1				
rs2917666	0.33	0.69	0.48	0.45
rs689453	0.91	0.53	0.55	0.52
rs1800566	0.90	0.64	0.84	0.34
rs10517	0.40	1.00	0.19	1.00

*HWE p value is not calculated

Table I2: Hardy-Weinberg equilibrium (HWE) p-values for SNPs in CYP1A1 and AhR genes for total populations and stratified by ethnicity

SNP ID	Total Population	Asian	Black	White	Others
AhR					
rs2074113	0.90	0.71	1.00	0.72	0.43
rs2066853	0.002	0.58	0.30	0.35	0.80
rs17722841	0.48	0.60	0.70	0.40	0.56
rs17779352	0.48	0.26	0.58	0.70	0.66
rs2282885	0.07	1.00	0.58	0.9	0.12
CYP1A1					
rs2606345	2.46E-11	0.46	0.14	1.00	0.58
?rs1799814	4.40E-04	0.07	0.82	0.02	0.27
rs17861115	0.004	0.13	0.44	0.36	0.05
rs2198843	1.90E-04	0.49	1.00	0.55	0.14

* SNP which is not in HWE ($p<0.05$).

Table I3: SNP linkage disequilibrium (*D' values >0.7 highlight in bold text*) for genes within the confirmatory and exploratory panel. Based on data from the 1000 genomes study: intergrated phase 1, version 3 (March 2012). Calculated using the linkage disequilibrium calculator, caprica.genetics.kcl.ac.uk.

Chromosome 11		GSTP1 *rs4147581*	GSTP1 *rs1138272*	GSTP1 *rs1871042*	GSTP1 *rs749174*	GSTP1 *rs1695*
GSTP1	*rs4147581*	-	**-1.000**	**-0.983**	**-0.983**	**-0.983**
GSTP1	*rs1138272*		-	**1.000**	**1.000**	**1.000**
GSTP1	*rs1871042*			-	**1.000**	**0.963**
GSTP1	*rs749174*				-	**0.969**
GSTP1	*rs1695*					-

Chromosome 16		NQO1 *rs2917666*	NQO1 *rs689453*	NQO1 *rs1800566*	NQO1 *rs10517*
NQO1	*rs2917666*	-	**1.000**	**-0.971**	**0.945**
NQO1	*rs689453*		-	**-1.000**	**1.000**
NQO1	*rs1800566*			-	**1.000**
NQO1	*rs10517*				-

Chromosome 7		Ahr *rs2074113*	Ahr *rs2066853*	Ahr *rs17722841*	Ahr *rs17779352*	Ahr *rs2282885*
Ahr	*rs2074113*	-	**1.000**	**-0.866**	**-1.000**	**-0.942**
Ahr	*rs2066853*		-	**-0.869**	**-1.000**	**-0.944**
Ahr	*rs17722841*			-	0.030	**0.987**
Ahr	*rs17779352*				-	-0.509
Ahr	*rs2282885*					-

Chromosome 15		CYP1A1 *rs2606345*	CYP1A1 *rs1799814*	CYP1A1 *rs17861115*	CYP1A1 *rs2472299*	CYP1A1 *rs2198843*
CYP1A1	*rs2606345*	-	**-1.000**	**-0.946**	0.170	**-0.973**
CYP1A1	*rs1799814*		-	**-1.000**	**1.000**	**1.000**
CYP1A1	*rs17861115*			-	0.659	**0.916**
CYP1A1	*rs2472299*				-	**0.775**
CYP1A1	*rs2198843*					-

Table I4: Unadjusted and adjusted beta-coefficients (β) and 95% confidence intervals (95% CI) for the association between GSTM1, GSTP1 and NQO1 genotypes and lung function

	FEV_1		FVC	
Genotype	**Crude (n=898)**	**Adjusted (n=866)**	**Crude (n=867)**	**Adjusted (n=837)**
GSTM1				
rs366631 (C/T)				
CT	Reference	Reference	Reference	Reference
TT	0.0526**[0.0166,0.0886]	0.0198[-0.0054,0.0451]	0.0552*[0.0128,0.0975]	0.0086[-0.0207,0.0379]
GSTP1				
rs749174 (Intron, 5 C/T)				
CC	Reference	Reference	Reference	Reference
CT	-0.0109[-0.0483,0.0265]	-0.0198[-0.0458,0.0063]	-0.0179[-0.0617,0.0260]	-0.0278[-0.0579,0.0022]
TT	0.0608[-0.0057,0.1273]	-0.0012[-0.0476,0.0451]	0.0497[-0.0286,0.1280]	-0.0197[-0.0731,0.0337]
CT + TT	0.0015[-0.0341,0.0371]	-0.0166[-0.0413,0.0080]	-0.0063[-0.0480,0.0354]	-0.0264[-0.0548,0.0020]
rs1695 (Exon 5, Ile105Val)				
Ile/Ile	Reference	Reference	Reference	Reference
Ile/Val	-0.0366[-0.0748,0.0017]	-0.0340*[-0.0604,-0.0075]	-0.0505*[-0.0953,-0.0058]	-0.0436**[-0.0741,-0.0132]
Val/Val	-0.0180[-0.0736,0.0376]	-0.0249[-0.0639,0.0141]	-0.0405[-0.1058,0.0248]	-0.0369[-0.0818,0.0080]
Ile/Val + Val/Val	-0.0321[-0.0680,0.0038]	-0.0319*[-0.0567,-0.0071]	-0.0480*[-0.0901,-0.0059]	-0.0422**[-0.0707,-0.0136]
NQO1				
rs2917666 (3´ UTR (C/G))				
CC	Reference	Reference	Reference	Reference
CG	-0.0150[-0.0577,0.0276]	-0.0081[-0.0384,0.0221]	-0.0184[-0.0686,0.0319]	-0.0039[-0.0390,0.0311]
GG	-0.0321[-0.0818,0.0175]	0.0108[-0.0254,0.0471]	-0.0360[-0.0944,0.0223]	0.0241[-0.0180,0.0661]
CG + GG	-0.0208[-0.0608,0.0192]	-0.0026[-0.0313,0.0261]	-0.0244[-0.0715,0.0228]	0.004[-0.0293,0.0374]
rs689453 (Exon 2, Glu24 (G/A))				
GG	Reference	Reference	Reference	Reference
GA	0.0157[-0.0511,0.0824]	-0.0207[-0.0675,0.0261]	0.0026[-0.0765,0.0817]	-0.0243[-0.0789,0.0304]
AA	-0.0622[-0.4388,0.3144]	0.0884[-0.1690,0.3457]	-0.2044[-0.6381,0.2292]	-0.0280[-0.3202,0.2641]
GA + AA	0.0136[-0.0522,0.0793]	-0.0174[-0.0632,0.0284]	-0.0035[-0.0814,0.0744]	-0.0244[-0.0778,0.0291]
rs1800566 (Exon 6, Pro187Ser)				
Pro/Pro	Reference	Reference	Reference	Reference
Pro/Ser	0.0137[-0.0242,0.0515]	0.0195[-0.0071,0.0461]	0.0186[-0.0257,0.0629]	0.0354*[0.0048,0.0660]
Ser/Ser	-0.0292[-0.1068,0.0485]	0.0104[-0.0440,0.0649]	-0.0350[-0.1264,0.0564]	0.0140[-0.0491,0.0772]
Pro/Ser + Ser/Ser	0.0077[-0.0285,0.0439]	0.0184[-0.0070,0.0438]	0.0113[-0.0312,0.0537]	0.0327*[0.0034,0.0620]
rs10517 (3´ UTR (C/T))				
CC	Reference	Reference	Reference	Reference
CT	-0.0035[-0.0428,0.0358]	0.0109[-0.0162,0.0380]	-0.0118[-0.0576,0.0341]	0.0145[-0.0166,0.0456]
TT	-0.0037[-0.1124,0.1049]	-0.0094[-0.0849,0.0661]	0.0064[-0.1240,0.1368]	-0.0207[-0.1082,0.0669]
CT + TT	-0.0036[-0.0416,0.0344]	0.0091[-0.0170,0.0352]	-0.0103[-0.0548,0.0341]	0.0116[-0.0184,0.0416]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. ***p<0.001, ** p<0.01, *p<0.05, #p<0.1 (interaction terms)

Table I5: Effects of GSTM1, GSTP1 and NQO1 genotypes on FEV_1 and acute pollutant exposures (24 hour)

		FEV_1			
Genotype	**Subjects (n)**	**NOx 20m β(95% CI)**	**NO_2 20m β(95% CI)**	**PM_{10} 20m β (95% CI)**	**$PM_{2.5}$ 20m β (95% CI)**
GSTM1					
rs366631 C/T					
CT	482	Reference	Reference	Reference	Reference
TT	383	45.9E-06[-0.0005,0.0006]	-21.4E-06[-0.0023,0.0022]	-0.0005[-0.0031,0.0021]	-0.0012[-0.0049,0.0025]
GSTP1					
rs749174 Intron 5 C/T					
CC	455	Reference	Reference	Reference	Reference
CT	340	30.6E-06[-0.0006,0.0006]	-0.0001[-0.0024,0.0023]	-0.0004[-0.0031,0.0023]	-0.0002[-0.0041,0.0036]
TT	71	-0.0002[-0.0012,0.0009]	-0.0004[-0.0044,0.0036]	-0.0005[-0.0058,0.0049]	-0.0021[-0.0095,0.0052]
CT + TT	411	2.40E-06[-0.0006,0.0006]	-0.0001[-0.0024,0.0021]	-0.0004[-0.0030,0.0022]	-0.0004[-0.0041,0.0033]
rs1695 Exon 5 Ile105Val					
Ile/Ile	379	Reference	Reference	Reference	Reference
Ile/Val	371	0.0005[-0.0001,0.0011]	0.0021#[-0.0004,0.0045]	0.0012[-0.0016,0.0040]	0.0018[-0.0021,0.0058]
Val/Val	115	0.0005[-0.0005,0.0014]	0.0015[-0.0020,0.0050]	0.0026[-0.0017,0.0070]	0.0035[-0.0027,0.0096]
Ile/Val + Val/Val	486	0.0005#[-0.0001,0.0010]	0.0019#[-0.0003,0.0042]	0.0015[-0.0011,0.0041]	0.0022[-0.0015,0.0059]
NQO1					
rs2917666 3´ UTR (C/G)					
CC	231	Reference	Reference	Reference	Reference
CG	424	0.0001[-0.0005,0.0008]	0.0001[-0.0026,0.0027]	0.0013[-0.0018,0.0044]	0.0021[-0.0023,0.0066]
GG	211	0.0003[-0.0005,0.0011]	0.0015[-0.0016,0.0047]	0.0013[-0.0024,0.0050]	0.0021[-0.0033,0.0074]
CG + GG	635	0.0002[-0.0004,0.0008]	0.0005[-0.0019,0.0030]	0.0013[-0.0016,0.0042]	0.0021[-0.0021,0.0063]
rs689453 Exon 2 Glu24 (G/A)					
GG	799	Reference	Reference	Reference	Reference
GA	65	-0.0001[-0.0011,0.0009]	-0.0007[-0.0046,0.0032]	0.0004[-0.0041,0.0048]	0.0012[-0.0050,0.0075]
AA	2	-0.0092[-0.0298,0.0115]	-0.0268[-0.0868,0.0331]	-0.0721[-0.2379,0.0938]	0.0750[-0.0981,0.2482]
GA + AA	67	-31.3E-06[-0.0010,0.0009]	-0.0006[-0.0044,0.0032]	0.0006[-0.0038,0.0049]	0.0016[-0.0046,0.0077]
rs1800566 Exon 6 Pro187Ser					
Pro/Pro	505	Reference	Reference	Reference	Reference
Pro/Ser	311	47.4E-06[-0.0005,0.0006]	0.0003[-0.0020,0.0027]	0.0006[-0.0021,0.0033]	0.0007[-0.0031,0.0046]
Ser/Ser	50	0.0002[-0.0016,0.0020]	-0.0006[-0.0061,0.0048]	0.0012[-0.0046,0.0070]	0.0051[-0.0037,0.0140]
Pro/Ser + Ser/Ser	361	0.0001[-0.0005,0.0006]	0.0002[-0.0021,0.0024]	0.0007[-0.0019,0.0032]	0.0012[-0.0025,0.0049]
rs10517 3´ UTR (C/T)					
CC	584	Reference	Reference	Reference	Reference
CT	257	-0.0001[-0.0006,0.0005]	-0.0004[-0.0029,0.0021]	0.0010[-0.0019,0.0038]	0.0004[-0.0036,0.0044]
TT	24	0.0005[-0.0017,0.0026]	0.0021[-0.0053,0.0096]	-0.0018[-0.0120,0.0083]	-0.0036[-0.0190,0.0118]
CT + TT	841	-43.4E-06[-0.0006,0.0005]	-0.0002[-0.0025,0.0022]	0.0009[-0.0018,0.0036]	0.0003[-0.0035,0.0042]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. ***p<0.001, ** p<0.01, *p<0.05, #p<0.1 (interaction terms)

Table I6: Effects of GSTM1, GSTP1 and NQO1 genotypes on FVC and acute pollutant exposures (24 hour)

		FVC			
Genotype	**Subjects (n)**	**NOx 20m β(95% CI)**	**NO_2 20m β(95% CI)**	**PM_{10} 20m β (95% CI)**	**$PM_{2.5}$ 20m β (95% CI)**
GSTM1					
rs366631 C/T					
CT	466	Reference	Reference	Reference	Reference
TT	370	0.0003[-0.0004,0.0009]	0.0008[-0.0019,0.0034]	-0.0001[-0.0032,0.0029]	-0.0002[-0.0045,0.0041]
GSTP1					
rs749174 Intron 5 C/T					
CC	438	Reference	Reference	Reference	Reference
CT	330	0.0001[-0.0006,0.0007]	-0.0001[-0.0028,0.0026]	0.0002[-0.0029,0.0034]	-0.0001[-0.0045,0.0044]
TT	69	-0.0004[-0.0016,0.0008]	-0.0024[-0.0071,0.0023]	-0.0002[-0.0063,0.0059]	-0.0022[-0.0106,0.0063]
CT + TT	399	0.6E-06[-0.0006,0.0006]	-0.0005[-0.0031,0.0021]	0.0002[-0.0028,0.0032]	-0.0003[-0.0046,0.0039]
rs1695 Exon 5 Ile105Val					
Ile/Ile	364	Reference	Reference	Reference	Reference
Ile/Val	361	0.0005[-0.0002,0.0012]	0.0019[-0.0008,0.0047]	0.0017[-0.0015,0.0049]	0.0025[-0.0021,0.0071]
Val/Val	111	0.0003[-0.0007,0.0014]	0.0004[-0.0037,0.0044]	0.0032[-0.0018,0.0082]	0.0043[-0.0028,0.0114]
Ile/Val + Val/Val	472	0.0005[-0.0002,0.0011]	0.0016[-0.0010,0.0042]	0.002[-0.0010,0.0050]	0.0029[-0.0014,0.0072]
NQO1					
rs2917666 3´ UTR (C/G)					
CC	220	Reference	Reference	Reference	Reference
CG	411	0.0002[-0.0006,0.0010]	0.0003[-0.0027,0.0033]	0.0018[-0.0018,0.0054]	0.0030[-0.0023,0.0082]
GG	206	0.0006[-0.0004,0.0015]	0.0018[-0.0018,0.0055]	0.0018[-0.0025,0.0061]	0.0035[-0.0027,0.0097]
CG + GG	617	0.0003[-0.0004,0.0010]	0.0008[-0.0020,0.0036]	0.0018[-0.0016,0.0052]	0.0031[-0.0018,0.0081]
rs689453 Exon 2 Glu24 (G/A)					
GG	774	Reference	Reference	Reference	Reference
GA	61	-0.0003[-0.0014,0.0009]	-0.0010[-0.0054,0.0035]	0.0001[-0.0050,0.0052]	0.0020[-0.0052,0.0092]
AA	2	-0.0055[-0.0290,0.0180]	-0.0160[-0.0842,0.0522]	-0.0436[-0.2321,0.1449]	0.0459[-0.1509,0.2427]
GA + AA	63	-0.0003[-0.0014,0.0009]	-0.001[-0.0053,0.0033]	0.0001[-0.0049,0.0050]	0.002[-0.0050,0.0091]
rs1800566 Exon 6 Pro187Ser					
Pro/Pro	487	Reference	Reference	Reference	Reference
Pro/Ser	302	-0.0001[-0.0007,0.0006]	-37.0E-06[-0.0028,0.0027]	0.0007[-0.0025,0.0038]	0.0008[-0.0037,0.0052]
Ser/Ser	48	0.0001[-0.0019,0.0022]	0.0002[-0.0059,0.0064]	0.0001[-0.0066,0.0068]	0.0032[-0.0069,0.0134]
Pro/Ser + Ser/Ser	350	-44.7E-06[-0.0007,0.0006]	-25.4E-06[-0.0026,0.0026]	0.0006[-0.0024,0.0036]	0.0011[-0.0032,0.0053]
rs10517 3´ UTR (C/T)					
CC	562	Reference	Reference	Reference	Reference
CT	252	0.0002[-0.0005,0.0008]	0.0005[-0.0023,0.0034]	0.0015[-0.0018,0.0047]	0.0012[-0.0034,0.0058]
TT	23	0.0002[-0.0023,0.0028]	0.0005[-0.0085,0.0095]	-0.0014[-0.0130,0.0101]	-0.0039[-0.0215,0.0138]
CT + TT	814	0.0002[-0.0005,0.0008]	0.0006[-0.0022,0.0033]	0.0014[-0.0017,0.0046]	0.0011[-0.0033,0.0056]

Table I7: Effects of GSTM1, GSTP1 and NQO1 genotypes on the association between FVC and sub-chronic pollutant exposures (7 day averages)

		FEV_1			
Genotype	**Subjects (n)**	**NOx 20m β(95% CI)**	**NO_2 20m β(95% CI)**	**PM_{10} 20m β (95% CI)**	**$PM_{2.5}$ 20m β (95% CI)**
GSTM1					
rs366631 C/T					
CT	482	Reference	Reference	Reference	Reference
TT	383	0.0006[-0.0010,0.0021]	0.0009[-0.0034,0.0052]	-0.0015[-0.0062,0.0031]	-0.0015[-0.0082,0.0053]
GSTP1					
rs749174 Intron 5 C/T					
CC	455	Reference	Reference	Reference	Reference
CT	340	-0.0005[-0.0020,0.0011]	-0.0019[-0.0062,0.0024]	-0.0027[-0.0074,0.0020]	-0.0033[-0.0103,0.0037]
TT	71	-0.0005[-0.0037,0.0027]	-0.0010[-0.0093,0.0072]	-0.0005[-0.0102,0.0093]	-0.0057[-0.0193,0.0078]
CT + TT	411	-0.0005[-0.0020,0.0010]	-0.0019[-0.0059,0.0022]	-0.0025[-0.0070,0.0020]	-0.0038[-0.0104,0.0029]
rs1695 Exon 5 Ile105Val					
Ile/Ile	379	Reference	Reference	Reference	Reference
Ile/Val	371	0.0010[-0.0005,0.0026]	0.0018[-0.0026,0.0062]	0.0002[-0.0047,0.0050]	0.0014[-0.0058,0.0086]
Val/Val	115	0.0004[-0.0020,0.0027]	0.0015[-0.0048,0.0078]	0.0009[-0.0060,0.0078]	-0.0009[-0.0109,0.0092]
Ile/Val + Val/Val	486	0.0009[-0.0006,0.0023]	0.0018[-0.0023,0.0058]	0.0004[-0.0040,0.0048]	0.0008[-0.0057,0.0074]
NQO1					
rs2917666 3´ UTR (C/G)					
CC	231	Reference	Reference	Reference	Reference
CG	424	0.0006[-0.0011,0.0023]	0.0013[-0.0035,0.0060]	0.0046#[-0.0006,0.0097]	0.0064#[-0.0011,0.0140]
GG	211	0.0006[-0.0014,0.0025]	0.0011[-0.0044,0.0066]	-0.0001[-0.0061,0.0059]	1.48E-06[-0.0090,0.0090]
CG + GG	635	0.0006[-0.0009,0.0021]	0.0012[-0.0031,0.0055]	0.0029[-0.0018,0.0075]	0.0042[-0.0027,0.0112]
rs689453 Exon 2 Glu24 (G/A)					
GG	799	Reference	Reference	Reference	Reference
GA	65	-0.0012[-0.0040,0.0017]	-0.0024[-0.0099,0.0051]	0.0015[-0.0077,0.0106]	0.0001[-0.0125,0.0127]
AA	2	-0.0546[-0.1784,0.0692]	-0.1405[-0.4569,0.1760]	0.0403[-0.0505,0.1312]	0.0371[-0.0473,0.1214]
GA + AA	67	-0.0011[-0.0039,0.0017]	-0.0022[-0.0097,0.0052]	0.0016[-0.0074,0.0106]	0.0009[-0.0115,0.0133]
rs1800566 Exon 6 Pro187Ser					
Pro/Pro	505	Reference	Reference	Reference	Reference
Pro/Ser	311	-0.0001[-0.0017,0.0015]	-0.0003[-0.0047,0.0042]	-0.0008[-0.0056,0.0040]	-0.0020[-0.0091,0.0050]
Ser/Ser	50	-0.0007[-0.0040,0.0026]	-0.0040[-0.0128,0.0049]	-0.0022[-0.0137,0.0093]	0.0011[-0.0156,0.0179]
Pro/Ser + Ser/Ser	361	-0.0002[-0.0017,0.0013]	-0.0009[-0.0051,0.0032]	-0.0009[-0.0055,0.0037]	-0.0016[-0.0084,0.0051]
rs10517 3´ UTR (C/T)					
CC	584	Reference	Reference	Reference	Reference
CT	257	0.0007[-0.0010,0.0023]	0.0015[-0.0031,0.0062]	0.0038[-0.0016,0.0092]	0.0052[-0.0026,0.0131]
TT	24	0.0009[-0.0042,0.0061]	0.0023[-0.0123,0.0168]	0.0089[-0.0033,0.0212]	0.0135[-0.0044,0.0314]
CT + TT	841	0.0006[-0.0010,0.0023]	0.0016[-0.0029,0.0061]	0.0043#[-0.0008,0.0093]	0.0059[-0.0014,0.0132]

Table I8: Effects of GSTM1, GSTP1 and NQO1 genotypes on the association between FVC and sub-chronic pollutant exposures (7 day averages)

		FVC			
Genotype	**Subjects (n)**	**NOx 20m β(95% CI)**	**NO_2 20m β(95% CI)**	**PM_{10} 20m β (95% CI)**	**$PM_{2.5}$ 20m β (95% CI)**
GSTM1					
rs366631 C/T					
CT	466	Reference	Reference	Reference	Reference
TT	370	0.0010[-0.0008,0.0028]	0.0019[-0.0031,0.0069]	-0.0013[-0.0068,0.0042]	-0.0007[-0.0087,0.0073]
GSTP1					
rs749174 Intron 5 C/T					
CC	438	Reference	Reference	Reference	Reference
CT	330	-0.0002[-0.0019,0.0016]	-0.0017[-0.0066,0.0033]	-0.0016[-0.0071,0.0039]	-0.0033[-0.0115,0.0050]
TT	69	-0.0018[-0.0056,0.0021]	-0.0056[-0.0157,0.0045]	-0.0039[-0.0151,0.0073]	-0.0091[-0.0247,0.0064]
CT + TT	399	-0.0004[-0.0021,0.0013]	-0.0022[-0.0069,0.0025]	-0.002[-0.0072,0.0033]	-0.0042[-0.0120,0.0036]
rs1695 Exon 5 Ile105Val					
Ile/Ile	364	Reference	Reference	Reference	Reference
Ile/Val	361	0.0008[-0.0010,0.0026]	0.0006[-0.0044,0.0056]	0.0020[-0.0038,0.0077]	0.0026[-0.0058,0.0111]
Val/Val	111	0.0006[-0.0021,0.0033]	0.0022[-0.0052,0.0096]	0.0001[-0.0078,0.0080]	-0.0022[-0.0137,0.0094]
Ile/Val + Val/Val	472	0.0008[-0.0009,0.0024]	0.001[-0.0037,0.0057]	0.0015[-0.0037,0.0067]	0.0014[-0.0063,0.0090]
NQO1					
rs2917666 3´ UTR (C/G)					
CC	220	Reference	Reference	Reference	Reference
CG	411	0.0013[-0.0007,0.0032]	0.0037[-0.0018,0.0091]	0.0083**[0.0024,0.0143]	0.0123**[0.0035,0.0211]
GG	206	0.0010[-0.0013,0.0032]	0.0021[-0.0042,0.0084]	0.0038[-0.0033,0.0109]	0.0075[-0.0032,0.0181]
CG + GG	617	0.0012[-0.0005,0.0029]	0.0032[-0.0017,0.0082]	0.0067*[0.0013,0.0121]	0.0107**[0.0026,0.0188]
rs689453 Exon 2 Glu24 (G/A)					
GG	774	Reference	Reference	Reference	Reference
GA	61	-0.0009[-0.0042,0.0024]	-0.0019[-0.0106,0.0068]	0.0028[-0.0076,0.0132]	0.0037[-0.0109,0.0182]
AA	2	-0.0323[-0.1732,0.1086]	-0.0839[-0.4444,0.2766]	0.0257[-0.0775,0.1290]	0.0243[-0.0715,0.1202]
GA + AA	63	-0.0009[-0.0042,0.0024]	-0.0018[-0.0104,0.0068]	0.0031[-0.0072,0.0133]	0.0042[-0.0101,0.0185]
rs1800566 Exon 6 Pro187Ser					
Pro/Pro	487	Reference	Reference	Reference	Reference
Pro/Ser	302	-0.0004[-0.0023,0.0014]	-0.0010[-0.0061,0.0042]	0.0007[-0.0049,0.0064]	-0.0002[-0.0085,0.0081]
Ser/Ser	48	-0.0003[-0.0040,0.0035]	-0.0027[-0.0129,0.0075]	0.0008[-0.0123,0.0140]	0.0070[-0.0124,0.0264]
Pro/Ser + Ser/Ser	350	-0.0004[-0.0022,0.0013]	-0.0013[-0.0061,0.0035]	0.0008[-0.0046,0.0062]	0.0007[-0.0072,0.0086]
rs10517 3´ UTR (C/T)					
CC	562	Reference	Reference	Reference	Reference
CT	252	0.0013[-0.0006,0.0032]	0.0030[-0.0024,0.0083]	0.0058#[-0.0004,0.0120]	0.0085#[-0.0006,0.0175]
TT	23	0.0006[-0.0057,0.0070]	-0.0002[-0.0184,0.0180]	0.0107[-0.0036,0.0250]	0.0152[-0.0059,0.0364]
CT + TT	814	0.0012[-0.0007,0.0030]	0.0027[-0.0025,0.0078]	0.0061*[0.0003,0.0118]	0.0086*[0.0001,0.0170]

Table I9: Unadjusted and adjusted beta-coefficients (β) and 95% confidence intervals (95% CI) for the association between AhR genotypes and lung function

Genotype	FEV1		FVC	
	Unadjusted (n=894)	Adjusted (n=862)	Unadjusted (n=864)	Adjusted (n=834)
rs2074113 C/A				
CC	Reference	Reference	Reference	Reference
CA	-0.0201[-0.0614,0.0212]	-0.0080[-0.0366,0.0206]	-0.0202[-0.0686,0.0283]	-0.0055[-0.0386,0.0275]
AA	-0.0605[-0.1790,0.0580]	-0.0089[-0.0913,0.0736]	-0.0700[-0.2068,0.0667]	-0.0040[-0.0976,0.0897]
CA+AA	-0.0236[-0.0636,0.0164]	-0.0083[-0.0359,0.0192]	-0.0246[-0.0715,0.0223]	-0.0059[-0.0377,0.0259]
rs2066853 Exon 10 Arg554Lys (G/A)				
Arg/Arg	Reference	Reference	Reference	Reference
Arg/Lys	-0.0250[-0.0638,0.0138]	-0.0069[-0.0347,0.0209]	-0.0368[-0.0824,0.0088]	-0.0066[-0.0388,0.0256]
Lys/Lys	-0.1084**[-0.1777,-0.0390]	-0.0471[-0.0972,0.0030]	-0.1213**[-0.2017,-0.0410]	-0.0391[-0.0963,0.0181]
Arg/Lys+Lys/Lys	-0.0398*[-0.0765,-0.0031]	-0.0131[-0.0397,0.0134]	-0.0522*[-0.0953,-0.0091]	-0.0119[-0.0426,0.0188]
rs17722841 G/A				
GG	Reference	Reference	Reference	Reference
GA	0.1028***[0.0504,0.1551]	0.0275[-0.0107,0.0658]	0.1149***[0.0533,0.1764]	0.0068[-0.0372,0.0509]
AA	0.0647[-0.1362,0.2656]	0.0355[-0.1026,0.1736]	0.0445[-0.1878,0.2767]	-0.0033[-0.1603,0.1537]
GA+AA	0.1012***[0.0502,0.1523]	0.0281[-0.0090,0.0653]	0.1116***[0.0516,0.1715]	0.0067[-0.0361,0.0495]
rs17779352 T/C				
TT	Reference	Reference	Reference	Reference
TC	0.0159[-0.0366,0.0684]	-0.0006[-0.0372,0.0359]	0.0130[-0.0488,0.0748]	0.0008[-0.0415,0.0432]
CC	0.1695[-0.0491,0.3880]	0.0397[-0.1423,0.2216]	0.0990[-0.1770,0.3749]	-0.0106[-0.2172,0.1960]
TC+CC	0.0229[-0.0285,0.0743]	0.0004[-0.0354,0.0361]	0.0164[-0.0443,0.0771]	0.0003[-0.0411,0.0417]
rs2282885 T/C				
TT	Reference	Reference	Reference	Reference
TC	0.0685***[0.0291,0.1078]	0.0313*[0.0023,0.0604]	0.0816***[0.0355,0.1277]	0.0340*[0.0004,0.0675]
CC	0.0381[-0.0389,0.1152]	0.0116[-0.0435,0.0668]	0.0564[-0.0349,0.1477]	-0.0042[-0.0686,0.0602]
TC+CC	0.0639***[0.0267,0.1011]	0.0285*[0.0008,0.0562]	0.0780***[0.0344,0.1217]	0.0288[-0.0032,0.0608]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.5, **p<0.01, ***p<0.001.

Table I10: Unadjusted and adjusted beta-coefficients (β) and 95% confidence intervals (95% CI) for the association between CYP1A1 genotypes and lung function

Genotype	FEV1		FVC	
	Unadjusted (n=888)	Adjusted (n=856)	Unadjusted (n=857)	Adjusted (n=827)
rs2606345 G/T				
GG	Reference	Reference	Reference	Reference
GT	ǂ0.0691***[0.0290,0.1091]	0.0144[-0.0166,0.0454]	0.0762**[0.0293,0.1231]	0.0090[-0.0267,0.0448]
TT	ǂ0.0836***[0.0339,0.1333]	-33.8e-06[-0.0409,0.0408]	0.1150***[0.0563,0.1737]	-0.0010[-0.0487,0.0466]
GT+TT	ǂ0.0740***[0.0378,0.1102]	0.0107[-0.0189,0.0402]	ǂ0.0890***[0.0464,0.1316]	0.0065[-0.0276,0.0407]
rs17861115 C/T				
CC	Reference	Reference	Reference	Reference
CT	-0.0529[-0.1074,0.0016]	-0.0122[-0.0506,0.0262]	-0.0687*[-0.1324,-0.0051]	-0.0147[-0.0591,0.0296]
TT	-0.0664[-0.2281,0.0953]	`	-0.1211[-0.3072,0.0650]	-0.0155[-0.1411,0.1101]
CT+TT	-0.0544*[-0.1066,-0.0022]	-0.0088[-0.0454,0.0278]	-0.0738*[-0.1347,-0.0129]	-0.0146[-0.0568,0.0275]
rs2198843 G/C				
GG	Reference	Reference	Reference	Reference
GC	-0.0466*[-0.0854,-0.0078]	-0.0125[-0.0406,0.0156]	ǂ-0.0636**[-0.1092,-0.0181]	-0.0146[-0.0471,0.0179]
CC	ǂ-0.0876***[-0.1395,-0.0358]	-0.0325[-0.0707,0.0058]	ǂ-0.1052***[-0.1655,-0.0448]	-0.0352[-0.0791,0.0088]
GC+CC	ǂ-0.0580**[-0.0941,-0.0220]	-0.0175[-0.0441,0.0092]	ǂ-0.0753***[-0.1176,-0.0330]	-0.0197[-0.0506,0.0111]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.5, **p<0.01, ***p<0.001. Blank cells indicate non-converged data.

Table I11: Effect modification of AhR genotypes on the association between FEV_1 and acute pollutant exposures (24 hour)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2074113 C/A					
CC	624	Reference	Reference	Reference	Reference
CA	218	-0.0003[-0.0010,0.0004]	-0.0008[-0.0035,0.0020]	-0.0011[-0.0043,0.0020]	-0.0010[-0.0054,0.0035]
AA	20	-0.0005[-0.0034,0.0024]	-0.0027[-0.0125,0.0072]	-0.0034[-0.0132,0.0065]	-0.0055[-0.0190,0.0081]
CA+AA	238	-0.0003[-0.0010,0.0003]	-0.0009[-0.0036,0.0017]	-0.0013[-0.0043,0.0017]	-0.0013[-0.0056,0.0029]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	516	Reference	Reference	Reference	Reference
Arg/Lys	281	0.0003[-0.0003,0.0009]	0.0020[-0.0005,0.0045]	0.0007[-0.0022,0.0036]	0.0009[-0.0031,0.0050]
Lys/Lys	68	-0.0006[-0.0021,0.0008]	-0.0021[-0.0067,0.0025]	-0.0019[-0.0072,0.0034]	-0.0032[-0.0112,0.0049]
Arg/Lys+Lys/Lys	349	0.0002[-0.0003,0.0008]	0.0013[-0.0010,0.0036]	0.0002[-0.0024,0.0029]	0.0003[-0.0035,0.0041]
rs17722841 G/A					
GG	742	Reference	Reference	Reference	Reference
GA	116	-0.0001[-0.0009,0.0008]	-0.0002[-0.0035,0.0031]	0.0013[-0.0026,0.0051]	0.0014[-0.0039,0.0068]
AA	7	0.0026*[0.0004,0.0048]	0.0119#[-0.0006,0.0245]	0.0130*[0.0014,0.0246]	0.0161*[0.0020,0.0302]
GA+AA	123	0.0002[-0.0006,0.0010]	0.0005[-0.0027,0.0037]	0.0022[-0.0014,0.0058]	0.0029[-0.0021,0.0079]
rs17779352 T/C					
TT	746	Reference	Reference	Reference	Reference
TC	115	-0.0004[-0.0012,0.0004]	-0.0013[-0.0045,0.0019]	-0.0013[-0.0052,0.0026]	-0.0022[-0.0076,0.0031]
CC	4	0.0016[-0.0005,0.0038]	0.0085[-0.0025,0.0195]	0.0117[-0.0026,0.0261]	0.0128[-0.0035,0.0291]
TC+CC	119	-0.0002[-0.0009,0.0006]	-0.0007[-0.0038,0.0025]	-0.0004[-0.0041,0.0033]	-0.0008[-0.0059,0.0043]
rs2282885 T/C					
TT	547	Reference	Reference	Reference	Reference
TC	264	-0.0002[-0.0008,0.0004]	0.0001[-0.0024,0.0026]	0.0014[-0.0016,0.0045]	0.0012[-0.0032,0.0056]
CC	51	-0.0003[-0.0011,0.0006]	-0.0010[-0.0053,0.0034]	0.0003[-0.0042,0.0048]	0.0003[-0.0058,0.0063]
TC+CC	315	-0.0002[-0.0008,0.0003]	-0.0001[-0.0024,0.0022]	0.001[-0.0017,0.0037]	0.0007[-0.0031,0.0045]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.5.

Table I12: Effect modification of AhR genotypes on the association between FVC and acute pollutant exposures (24 hour)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2074113 C/A					
CC	604	Reference	Reference	Reference	Reference
CA	210	-0.0001[-0.0009,0.0006]	48.5E06[-0.0031,0.0032]	0.0006[-0.0031,0.0042]	0.0005[-0.0046,0.0056]
AA	20	-0.0011[-0.0044,0.0022]	-0.0003[-0.0115,0.0108]	-0.0002[-0.0114,0.0110]	-0.0011[-0.0165,0.0142]
CA+AA	230	-0.0002[-0.0010,0.0006]	-8.75E-06[-0.0031,0.0031]	0.0005[-0.0030,0.0040]	0.0004[-0.0046,0.0053]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	498	Reference	Reference	Reference	Reference
Arg/Lys	270	0.0006[-0.0001,0.0013]	0.0031*[0.0002,0.0059]	0.0027[-0.0006,0.0060]	0.0032[-0.0014,0.0079]
Lys/Lys	68	-0.0007[-0.0024,0.0009]	-0.0021[-0.0073,0.0031]	0.0003[-0.0058,0.0063]	0.0003[-0.0089,0.0095]
Arg/Lys+Lys/Lys	338	0.0004[-0.0002,0.0011]	0.0022[-0.0005,0.0048]	0.0023[-0.0008,0.0053]	0.0028[-0.0016,0.0072]
rs17722841 G/A					
GG	717	Reference	Reference	Reference	Reference
GA	113	0.0001[-0.0009,0.0011]	-46.1E-06[-0.0039,0.0038]	0.0004[-0.0041,0.0049]	0.0008[-0.0054,0.0071]
AA	7	0.0013[-0.0012,0.0038]	0.0054[-0.0089,0.0198]	0.0075[-0.0057,0.0208]	0.0099[-0.0062,0.0260]
GA+AA	120	0.0002[-0.0007,0.0011]	0.0002[-0.0034,0.0039]	0.0009[-0.0033,0.0051]	0.0017[-0.0041,0.0074]
rs17779352 T/C					
TT	722	Reference	Reference	Reference	Reference
TC	110	-0.0003[-0.0012,0.0006]	-0.0009[-0.0046,0.0029]	-0.0012[-0.0056,0.0033]	-0.0014[-0.0076,0.0047]
CC	4	0.0011[-0.0013,0.0036]	0.0057[-0.0069,0.0182]	0.0078[-0.0085,0.0241]	0.0094[-0.0092,0.0279]
TC+CC	114	-0.0001[-0.0010,0.0007]	-0.0004[-0.0040,0.0032]	-0.0007[-0.0049,0.0036]	-0.0005[-0.0064,0.0053]
rs2282885 T/C					
TT	528	Reference	Reference	Reference	Reference
TC	257	-0.0001[-0.0008,0.0006]	0.0007[-0.0022,0.0035]	0.0011[-0.0023,0.0046]	0.0004[-0.0046,0.0055]
CC	48	-0.0001[-0.0013,0.0010]	-0.0013[-0.0065,0.0040]	0.0003[-0.0052,0.0059]	0.0002[-0.0073,0.0077]
TC+CC	305	-0.0001[-0.0008,0.0005]	0.0002[-0.0024,0.0029]	0.0008[-0.0024,0.0039]	29.2E-06[-0.0044,0.0045]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.5.

Table I13: Effect modification of AhR genotypes on the association between FEV_1 and sub-chronic pollutant exposures (7 days averages)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2074113 C/A					
CC	624	Reference	Reference	Reference	Reference
CA	218	0.0002[-0.0013,0.0018]	0.0012[-0.0031,0.0056]	-0.0008[-0.0056,0.0039]	-0.0015[-0.0086,0.0055]
AA	20	0.0027[-0.0026,0.0079]	0.0091[-0.0043,0.0225]	0.0037[-0.0151,0.0225]	0.0016[-0.0245,0.0277]
CA+AA	238		0.0017[-0.0025,0.0059]	-0.0006[-0.0053,0.0040]	-0.0014[-0.0083,0.0055]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	516	Reference	Reference	Reference	Reference
Arg/Lys	281	0.0014#[-0.0002,0.0030]		-0.0018[-0.0064,0.0029]	-0.0036[-0.0106,0.0033]
Lys/Lys	68	0.0002[-0.0020,0.0024]		0.0014[-0.0076,0.0105]	-0.0017[-0.0145,0.0111]
Arg/Lys+Lys/Lys	349		0.0029[-0.0012,0.0069]	-0.0015[-0.0059,0.0029]	-0.0035[-0.0101,0.0030]
rs17722841 G/A					
GG	742	Reference	Reference	Reference	Reference
GA	116	-0.0025#[-0.0052,0.0001]	-0.0050[-0.0119,0.0019]	-0.0027[-0.0111,0.0057]	-0.0078[-0.0196,0.0041]
AA	7	0.0048[-0.0023,0.0120]	0.0099[-0.0149,0.0347]	0.0146[-0.0149,0.0441]	0.0234[-0.0155,0.0623]
GA+AA	123	-0.0016[-0.0041,0.0009]	-0.0039[-0.0105,0.0028]	-0.0015[-0.0096,0.0066]	-0.0052[-0.0164,0.0060]
rs17779352 T/C					
TT	746	Reference	Reference	Reference	Reference
TC	115	0.0002[-0.0021,0.0024]		0.0014[-0.0047,0.0074]	0.0009[-0.0080,0.0098]
CC	4	0.0093[-0.0024,0.0210]		0.0756[-0.0189,0.1702]	0.0633[-0.0160,0.1426]
TC+CC	119	0.0004[-0.0018,0.0026]	0.0017[-0.0044,0.0078]	0.0015[-0.0045,0.0074]	0.0013[-0.0075,0.0100]
rs2282885 T/C					
TT	547	Reference	Reference	Reference	Reference
TC	264	-0.0002[-0.0019,0.0015]		0.0010[-0.0040,0.0060]	-0.0001[-0.0074,0.0072]
CC	51	0.0001[-0.0032,0.0034]		0.0056[-0.0068,0.0180]	0.0092[-0.0079,0.0264]
TC+CC	315	-0.0002[-0.0018,0.0014]	0.0005[-0.0039,0.0049]	0.0014[-0.0033,0.0062]	0.0009[-0.0060,0.0078]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1. Blank cells indicate non-converged data.

Table I14: Effect modification of AhR genotypes on the association between FVC and sub-chronic pollutant exposures (7 days averages)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2074113 C/A					
CC	604	Reference	Reference	Reference	Reference
CA	210	0.0004[-0.0014,0.0022]	0.0017[-0.0033,0.0068]	0.0019[-0.0037,0.0074]	0.0031[-0.0052,0.0114]
AA	20	0.0012[-0.0047,0.0072]	0.0084[-0.0068,0.0236]	0.0051[-0.0163,0.0264]	-0.0014[-0.0310,0.0283]
CA+AA	230	0.0004[-0.0013,0.0022]	0.0021[-0.0028,0.0070]	0.002[-0.0034,0.0074]	0.0029[-0.0052,0.0109]
rs2066853 Exon 10 Arg554Lys (G/A)					
Arg/Arg	498	Reference	Reference	Reference	Reference
Arg/Lys	270	0.0015[-0.0003,0.0034]	0.0043[-0.0009,0.0095]	0.0008[-0.0047,0.0062]	0.0005[-0.0077,0.0087]
Lys/Lys	68	-0.0001[-0.0026,0.0024]	0.0002[-0.0068,0.0072]	0.0029[-0.0074,0.0132]	-0.0003[-0.0150,0.0143]
Arg/Lys+Lys/Lys	338	0.0011[-0.0006,0.0027]	0.0031[-0.0016,0.0078]	0.0009[-0.0043,0.0061]	0.0001[-0.0075,0.0078]
rs17722841 G/A					
GG	717	Reference	Reference	Reference	Reference
GA	113	-0.0014[-0.0045,0.0016]	-0.0031[-0.0111,0.0049]	-0.0035[-0.0133,0.0064]	-0.0080[-0.0218,0.0058]
AA	7	0.0021[-0.0061,0.0102]	0.0017[-0.0266,0.0300]	0.0090[-0.0246,0.0425]	0.0161[-0.0281,0.0604]
GA+AA	120	-0.001[-0.0039,0.0019]	-0.0025[-0.0102,0.0051]	-0.0027[-0.0120,0.0067]	-0.0061[-0.0191,0.0069]
rs17779352 T/C					
TT	722	Reference	Reference	Reference	Reference
TC	110	0.0008[-0.0018,0.0034]	0.0029[-0.0045,0.0103]	0.0005[-0.0067,0.0077]	0.0020[-0.0087,0.0127]
CC	4	0.0066[-0.0067,0.0199]	0.0180[-0.0270,0.0630]	0.0500[-0.0575,0.1575]	0.0446[-0.0454,0.1347]
TC+CC	114	0.001[-0.0015,0.0035]	0.0033[-0.0039,0.0104]	0.0007[-0.0064,0.0078]	0.0025[-0.0080,0.0130]
rs2282885 T/C					
TT	528	Reference	Reference	Reference	Reference
TC	257	0.0003[-0.0017,0.0022]	0.0021[-0.0034,0.0075]	0.0005[-0.0053,0.0062]	-0.0012[-0.0096,0.0072]
CC	48	-0.0001[-0.0040,0.0039]	-0.0036[-0.0148,0.0075]	0.0038[-0.0113,0.0188]	0.0050[-0.0156,0.0256]
TC+CC	305	0.0002[-0.0016,0.0020]	0.0012[-0.0039,0.0062]	0.0007[-0.0048,0.0062]	-0.0008[-0.0088,0.0073]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.

Table I15: Effect modification of CYP1A1 genotypes on the association between FEV_1 and acute pollutant exposures (24 hour)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2606345 G/T					
GG	391	Reference	Reference	Reference	Reference
GT	319	-0.0003[-0.0010,0.0003]	-0.0020[-0.0044,0.0005]	-0.0014[-0.0044,0.0015]	-0.0008[-0.0050,0.0035]
TT	156	-0.0006#[-0.0014,0.0001]	-0.0029#[-0.0060,0.0002]	-0.0026[-0.0061,0.0008]	-0.0030[-0.0078,0.0018]
GT+TT	475	-0.0004[-0.0010,0.0001]	-0.0023*[-0.0045,-16.1E-06]	-0.0019[-0.0045,0.0008]	-0.0017[-0.0054,0.0021]
rs17861115 C/T					
CC	749	Reference	Reference	Reference	Reference
CT	106	-0.0003[-0.0011,0.0006]	-0.0007[-0.0041,0.0027]	-0.0028[-0.0070,0.0013]	-0.0019[-0.0081,0.0042]
TT	11	-0.0080**[-0.0139,-0.0020]	-0.0172#[-0.0347,0.0003]	-0.0161#[-0.0348,0.0027]	-0.0365*[-0.0683,-0.0047]
CT+TT	117	-0.0004[-0.0012,0.0004]		-0.0035#[-0.0075,0.0006]	-0.0032[-0.0092,0.0027]
rs2198843 G/C					
GG	373	Reference	Reference	Reference	Reference
GC	349	6.94E-06[-0.0006,0.0006]	0.0003[-0.0021,0.0027]	-0.0003[-0.0031,0.0025]	-0.0001[-0.0041,0.0040]
CC	144	0.0013*[0.0003,0.0023]	0.0048**[0.0012,0.0085]	0.0035#[-0.0006,0.0075]	0.0042[-0.0014,0.0098]
GC+CC	493	0.0002[-0.0004,0.0008]	0.0013[-0.0009,0.0036]	0.0006[-0.0020,0.0032]	0.0011[-0.0027,0.0048]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01. Blank cells indicate non-converged data.

.

Table I16: Effect modification of CYP1A1 genotypes on the association between FVC and acute pollutant exposures (24 hour)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2606345 G/T					
GG	380	Reference	Reference	Reference	Reference
GT	310	-0.0007#[-0.0014,37.6E-06]	-0.0033*[-0.0062,-0.0005]	-0.0034#[-0.0068,19.4E-06]	-0.0033[-0.0082,0.0016]
TT	147	-0.0006[-0.0015,0.0003]	-0.0028[-0.0064,0.0008]	-0.0027[-0.0068,0.0014]	-0.0037[-0.0094,0.0019]
GT+TT	457	-0.0007*[-0.0013,-9.29E-06]	-0.0032*[-0.0057,-0.0006]	-0.0031*[-0.0061,-0.0001]	-0.0035[-0.0078,0.0009]
rs17861115 C/T					
CC	723	Reference	Reference	Reference	Reference
CT	103	-49.2E-06[-0.0010,0.0009]	-0.0002[-0.0041,0.0038]	-0.0012[-0.0060,0.0036]	0.0003[-0.0068,0.0073]
TT	11	-0.0080*[-0.0148,-0.0012]	-0.0155[-0.0354,0.0044]	-0.0100[-0.0313,0.0114]	-0.0298[-0.0660,0.0064]
CT+TT	114	-0.0002[-0.0011,0.0008]	-0.0007[-0.0045,0.0032]	-0.0016[-0.0062,0.0030]	-0.0007[-0.0076,0.0061]
rs2198843 G/C					
GG	354	Reference	Reference	Reference	Reference
GC	340	-34.1E-06[-0.0007,0.0006]	-0.0005[-0.0032,0.0023]	-0.0005[-0.0037,0.0028]	0.0007[-0.0040,0.0053]
CC	143	₸0.0018**[0.0006,0.0030]	₸0.0066**[0.0024,0.0107]	0.0055*[0.0008,0.0101]	0.0067*[0.0003,0.0131]
GC+CC	483	0.0003[-0.0004,0.0009]	0.0011[-0.0015,0.0037]	0.0010[-0.0020,0.0040]	0.0023[-0.0020,0.0066]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01, ***p<0.001. ₸ Significant after the multiple testing procedures has been applied.

Table I17: Effect modification of CYP1A1 genotypes on the association between FEV_1 and sub-chronic pollutant exposures (7 day averages)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2606345 G/T					
GG	391	Reference	Reference	Reference	Reference
GT	319	-0.0023**[-0.0039,-0.0008]		-0.0070**[-0.0117,-0.0023]	-0.0086*[-0.0157,-0.0015]
TT	156	0.0001[-0.0022,0.0023]		0.0026[-0.0056,0.0108]	0.0029[-0.0081,0.0139]
GT+TT	475	-0.0018*[-0.0033,-0.0003]	-0.0051*[-0.0091,-0.0010]	-0.0052*[-0.0097,-0.0008]	-0.0061#[-0.0126,0.0004]
rs17861115 C/T					
CC	749	Reference	Reference	Reference	Reference
CT	106	0.0003[-0.0019,0.0026]		0.0001[-0.0072,0.0074]	0.0034[-0.0071,0.0139]
TT	11	-0.0017[-0.0108,0.0074]		-0.0090[-0.0595,0.0414]	-0.0093[-0.0723,0.0536]
CT+TT	117	0.0002[-0.0020,0.0024]	0.0008[-0.0051,0.0067]	-0.0001[-0.0073,0.0071]	0.003[-0.0073,0.0133]
rs2198843 G/C					
GG	373	Reference		Reference	Reference
GC	349	0.0008[-0.0008,0.0023]		0.0018[-0.0032,0.0068]	0.003[-0.0044,0.0104]
CC	144	0.0037**[0.0014,0.0061]		0.0055#[-0.0008,0.0118]	0.0071[-0.0020,0.0162]
GC+CC	493	0.0014#[-0.0000,0.0029]		0.0029[-0.0016,0.0074]	0.0042[-0.0024,0.0108]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01. Blank cells indicate non-converged data.

Table I18: Effect modification of CYP1A1 genotypes on the association between FVC and sub-chronic pollutant exposures (7 day averages)

Genotype	Subjects (n)	NOx β (95% CI)	NO_2 β (95% CI)	PM_{10} β (95% CI)	$PM_{2.5}$ β (95% CI)
rs2606345 G/T					
GG	380	Reference	Reference	Reference	Reference
GT	310	₸-0.0029**[-0.0047,-0.0012]	₸-0.0085***[-0.0135,-0.0035]	-0.0092**[-0.0148,-0.0036]	-0.0119**[-0.0203,-0.0035]
TT	147	-0.0007[-0.0033,0.0020]	-0.0026[-0.0100,0.0047]	0.0011[-0.0084,0.0107]	0.0016[-0.0113,0.0144]
GT+TT	457	₸-0.0024**[-0.0041,-0.0008]	₸-0.0072**[-0.0118,-0.0025]	-0.0072**[-0.0124,-0.0020]	-0.0088*[-0.0164,-0.0011]
rs17861115 C/T					
CC	723	Reference	Reference	Reference	Reference
CT	103	0.0009[-0.0017,0.0035]	0.0022[-0.0051,0.0094]	0.0045[-0.0047,0.0137]	0.0109[-0.0023,0.0240]
TT	11	-0.0032[-0.0136,0.0072]	-0.0087[-0.0344,0.0170]	-0.0143[-0.0717,0.0430]	-0.0275[-0.0990,0.0439]
CT+TT	114	0.0007[-0.0018,0.0032]	0.0014[-0.0055,0.0084]	0.004[-0.0050,0.0130]	0.0097[-0.0031,0.0226]
rs2198843 G/C					
GG	354	Reference	Reference	Reference	Reference
GC	340	0.0007[-0.0010,0.0025]	0.0035[-0.0016,0.0086]	0.0052#[-0.0007,0.0110]	0.0075#[-0.0013,0.0162]
CC	143	₸0.0052***[0.0026,0.0079]	₸0.0140***[0.0071,0.0210]	0.0094*[0.0022,0.0166]	0.0120*[0.0016,0.0224]
GC+CC	483	0.0018*[0.0001,0.0035]	0.0063**[0.0016,0.0111]	0.0065*[0.0012,0.0117]	0.0088*[0.0011,0.0166]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.#p<0.1, *p<0.05, **p<0.01, ***p<0.001. ₸ Significant after the multiple testing procedures has been applied.

Appendix J

Supplementary Data for Chapter 5

Table J1: Hardy-Weinberg equilibrium (HWE) p-values for SNPs in Nrf2, GCLM and SOD3 genes for the total populations and stratified by ethnicity

Genes/SNP ID	Total Population	Asian	Black	White	Others
		Nrf2			
rs2364723	3.00E-04	0.06	0.44	0.88	0.84
rs7557529	0.20	0.50	1.00	0.33	0.05
rs2001350	0.43	0.61	0.78	0.33	0.25
		GCLM			
rs2301022	0.02	0.13	0.52	0.18	0.16
rs3170633	2.96E-08	0.37	0.45	1.00	0.18
		SOD3			
ŦrsŦ2284659	1.45E-08	0.39	0.001	0.52	0.57
rs8192287	0.54	0.20	0.90	0.05	1.00
rs699473	0.004	0.52	0.74	0.51	0.57
rs13306703	0.07	1.00	0.73	2.99E-04	0.30
rs8192288	0.55	0.20	0.90	0.06	1.00

Table K1 shows the Hardy-Weinberg equilibrium p-values for SNPs Nrf2, GCLM and SOD3 genes for total population and stratified by ethnicity. Overall, all the SNPs have p-value >0.05 except for one SNP in SOD3, rs2284659 which was not in HW for either in total population or by ethnicity. Therefore, this SNP was excluded from further analysis.

Table J2: Linkage disequilibrium (*D' values >0.7 highlight in bold text* for selected SNPs within Nrf2, GCLM and SOD3. Based on data from the 1000 genomes study: intergrated phase 1, version 3 (March 2012). Calculated using the linkage disequilibrium calculator, caprica.genetics.kcl.ac.uk.

Chromosome 2		Nrf2 *rs2364723*	Nrf2 *rs7557529*	Nrf2 *rs2001350*
Nrf2	*rs2364723*	-	**0.983**	**0.960**
Nrf2	*rs7557529*		-	**0.951**
Nrf2	*rs2001350*			-

Chromosome 1		GCLM *rs2301022*	GCLM *rs3170633*
GCLM	*rs2301022*	-	0.661
GCLM	*rs3170633*		-

Chromosome 4		EC-SOD *rs2284659*	EC-SOD *rs8192287*	EC-SOD *rs699473*	EC-SOD *rs8192290*	EC-SOD *rs13306703*	EC-SOD *rs8192288*
EC-SOD	*rs2284659*	-	**-1.000**	**0.988**	**-0.945**	**-1.000**	**-1.000**
EC-SOD	*rs8192287*		-	**-1.000**	**-1.000**	**0.972**	**-1.000**
EC-SOD	*rs699473*			-	**-0.946**	**-1.000**	**-1.000**
EC-SOD	*rs8192290*				-	**0.885**	**-1.000**
EC-SOD	*rs13306703*					-	**0.972**
EC-SOD	*rs8192288*						-

Table J3: Distribution of genotypes and allele frequencies in Nrf2 and GLCM

Genotype/Allele Frequencies	Ethnicity								Total Population	
	Asian		Black		White		Others			
	n	%	n	%	n	%	n	%	n	%
Nrf2										
rs2364723 G/C										
GG	77	0.22	154	0.62	137	0.52	45	0.39	413	0.42
GC	154	0.44	85	0.34	107	0.40	55	0.48	401	0.41
CC	116	0.33	8	0.03	22	0.08	14	0.12	160	0.16
Allele G	308	0.44	393	0.80	381	0.72	145	0.64	1227	0.63
Alelle C	386	0.56	101	0.20	151	0.28	83	0.36	721	0.37
rs7557529 T/C										
TT	187	0.54	85	0.35	70	0.26	39	0.34	381	0.39
TC	131	0.38	119	0.49	125	0.47	64	0.56	439	0.45
CC	28	0.08	41	0.17	71	0.27	11	0.10	151	0.16
Allele T	505	0.73	289	0.59	265	0.50	142	0.62	1201	0.62
Allele C	187	0.27	201	0.41	267	0.50	86	0.38	741	0.38
rs2001350 A/G										
AA	268	0.77	187	0.76	209	0.79	92	0.81	756	0.78
AG	72	0.21	57	0.23	56	0.21	22	0.19	207	0.21
GG	6	0.02	3	0.01	1	0.00	0	0.00	10	0.01
Allele A	608	0.88	431	0.87	474	0.89	206	0.90	1719	0.88
Allele G	84	0.12	63	0.13	58	0.11	22	0.10	227	0.12
GCLM										
rs2301022 G/A										
GG	169	0.49	57	0.23	127	0.48	48	0.42	401	0.41
GA	136	0.39	118	0.48	120	0.45	46	0.40	420	0.43
AA	40	0.12	72	0.29	18	0.07	20	0.18	150	0.15
Allele G	474	0.69	232	0.47	374	0.71	142	0.62	1222	0.63
Allele A	216	0.31	262	0.53	156	0.29	86	0.38	720	0.37
rs3170633 G/A										
GG	227	0.65	25	0.10	107	0.40	44	0.39	403	0.41
GA	111	0.32	98	0.40	123	0.46	48	0.42	380	0.39
AA	9	0.03	124	0.50	35	0.13	22	0.19	190	0.20
Allele G	565	0.81	148	0.30	337	0.64	136	0.60	1186	0.61
Allele A	129	0.19	346	0.70	193	0.36	92	0.40	760	0.39

Minor allele for this ethnicity is different with other ethnicities for the particular SNP.
Genotype distribution for this ethnicity are not similar with other ethnicity group despite sharing same major/minor allele.
Similar genotype/allele frequencies within the block for the particular SNP but not in all ethnicity.
Similar genotype/allele frequencies within the block for the particular SNP but not in all ethnicity.

Table J4: Distribution of genotypes and allele frequencies in extra-cellular superoxide dismutase within the study population

Genotype/Allele Frequencies	Ethnicity								Total Population	
	Asian		Black		White		Others			
	n	%	n	%	n	%	n	%	n	%
rs8192287 G/T										
GG	247	0.72	241	0.98	230	0.86	92	0.81	810	0.84
GT	94	0.27	4	0.02	32	0.12	21	0.18	151	0.16
TT	4	0.01	0	0.00	4	0.02	1	0.01	9	0.01
Allele G	588	0.85	486	0.99	492	0.92	205	0.90	1771	0.91
Allele T	102	0.15	4	0.01	40	0.08	23	0.10	169	0.09
rs699473 C/T										
CC	91	0.26	139	0.56	40	0.15	30	0.26	300	0.31
CT	167	0.48	91	0.37	119	0.45	61	0.54	438	0.45
TT	88	0.25	17	0.07	106	0.40	23	0.20	234	0.24
Allele C	349	0.50	369	0.75	199	0.38	121	0.53	1038	0.53
Allele T	343	0.50	125	0.25	331	0.62	107	0.47	906	0.47
rs13306703 C/T										
CC	213	0.62	141	0.57	217	0.82	82	0.73	653	0.67
CT	117	0.34	93	0.38	39	0.15	27	0.24	276	0.28
TT	15	0.04	13	0.05	10	0.04	4	0.04	42	0.04
Allele C	543	0.79	375	0.76	473	0.89	191	0.85	1582	0.81
Allele T	147	0.21	119	0.24	59	0.11	35	0.15	360	0.19
rs8192288 G/T										
GG	249	0.72	243	0.98	229	0.86	91	0.80	812	0.83
GT	94	0.27	4	0.02	33	0.12	22	0.19	153	0.16
TT	4	0.01	0	0.00	4	0.02	1	0.01	9	0.01
Allele G	592	0.85	490	0.99	491	0.92	204	0.89	1777	0.91
Allele T	102	0.15	4	0.01	41	0.08	24	0.11	171	0.09

Minor allele for this ethnicity is different with other ethnicities for the particular SNP.

Table J5: Effect modification of Nrf2 and GCLM Genotypes on the association between 8-isoprostane and acute pollutant exposure (24 hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	394	Reference	Reference	Reference	Reference
GC	382	1.0017**[1.0005,1.0029]	1.0058*[1.0008,1.0109]	1.0072*[1.0015,1.0130]	1.0092*[1.0010,1.0175]
CC	155	1.0004[0.9985,1.0022]	1.0029[0.9961,1.0098]	1.0065[0.9982,1.0148]	1.0082[0.9959,1.0206]
GC+CC	537	1.0014*[1.0002,1.0025]	1.0050*[1.0003,1.0096]	1.0071**[1.0017,1.0125]	1.0090*[1.0013,1.0167]
rs7557529 T/C					
TT	358	Reference	Reference	Reference	Reference
TC	426	1.001[0.9996,1.0023]	1.003[0.9979,1.0081]	1.0027[0.9968,1.0087]	1.0043[0.9956,1.0131]
CC	145	0.9993[0.9976,1.0009]	0.9975[0.9909,1.0042]	0.9981[0.9904,1.0058]	0.9971[0.9860,1.0082]
TC+CC	571	1.0005[0.9992,1.0018]	1.0015[0.9967,1.0063]	1.0014[0.9958,1.0071]	1.0023[0.9941,1.0106]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	0.9992[0.9979,1.0005]	0.9975[0.9921,1.0030]	0.9974[0.9913,1.0036]	0.9938[0.9851,1.0026]
GG	10	1.0012[0.9931,1.0093]	1.001[0.9741,1.0287]	1.0084[0.9721,1.0460]	1.0226[0.9603,1.0889]
AG+GG	210	0.9993[0.9980,1.0005]	0.9977[0.9924,1.0031]	0.9977[0.9916,1.0038]	0.9943[0.9857,1.0030]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	403	1.0009[0.9997,1.0021]	1.0019[0.9969,1.0069]	1.0043[0.9986,1.0099]	1.0062[0.9981,1.0144]
AA	141	0.9994[0.9974,1.0013]	0.9983[0.9913,1.0054]	0.9975[0.9890,1.0061]	0.9942[0.9815,1.0070]
GA+AA	544	1.0006[0.9995,1.0018]	1.001[0.9964,1.0057]	1.0028[0.9974,1.0082]	1.0038[0.9961,1.0115]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360	0.9991[0.9978,1.0004]	0.9965[0.9914,1.0017]	0.9933*[0.9873,0.9992]	0.9913*[0.9827,0.9999]
AA	184	1.0009[0.9994,1.0024]	1.0026[0.9965,1.0089]	0.9953[0.9883,1.0023]	0.9948[0.9847,1.0050]
GA+AA	544	0.9998[0.9986,1.0010]	0.9987[0.9940,1.0034]	0.9939*[0.9885,0.9993]	0.9925[0.9847,1.0003]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.05, **p<0.01. Ŧ Significant after the multiple testing procedures has been applied.

Table J6: Effect modification of Nrf2 and GCLM Genotypes on the association between 8-oxodG and acute pollutant exposure (24 hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	374	Reference	Reference	Reference	Reference
GC	366	1.0018#[0.9998,1.0039]	1.0109*[1.0024,1.0194]	1.0129**[1.0032,1.0226]	1.0168*[1.0028,1.0310]
CC	145	1.001[0.9980,1.0041]	1.0042[0.9927,1.0158]	1.0063[0.9925,1.0203]	1.0111[0.9900,1.0327]
GC+CC	511	1.0016#[0.9997,1.0035]	1.0090*[1.0012,1.0168]	1.0113*[1.0023,1.0204]	1.0156*[1.0024,1.0290]
rs7557529 T/C					
TT	340	Reference	Reference	Reference	Reference
TC	406	1.0016[0.9994,1.0038]	1.0096*[1.0010,1.0182]	1.0118*[1.0018,1.0218]	1.0135#[0.9987,1.0286]
CC	138	0.9983[0.9955,1.0010]	0.9918[0.9808,1.0029]	0.9939[0.9812,1.0067]	0.9881[0.9698,1.0067]
TC+CC	544	1.0007[0.9986,1.0028]	1.0046[0.9965,1.0127]	1.0065[0.9971,1.0160]	1.0061[0.9920,1.0203]
rs2001350 A/G					
AA	684	Reference	Reference	Reference	Reference
AG	191	1.0012[0.9990,1.0033]	1.0085#[0.9994,1.0176]	1.0087#[0.9985,1.0191]	1.01[0.9953,1.0249]
GG	10	1.0036[0.9904,1.0170]	1.0148[0.9706,1.0609]	0.9882[0.9312,1.0487]	0.9879[0.8925,1.0936]
AG+GG	201	1.001[0.9989,1.0032]	1.0077#[0.9988,1.0167]	1.0075[0.9973,1.0178]	1.0086[0.9939,1.0234]
			GCLM		
rs2301022 G/A					
GG	360	Reference	Reference	Reference	Reference
GA	385	1.0003[0.9983,1.0023]	1.0013[0.9930,1.0097]	1.0026[0.9932,1.0121]	1.0047[0.9910,1.0186]
AA	137	0.9995[0.9963,1.0027]	1.0001[0.9884,1.0120]	1.0029[0.9887,1.0173]	1.0058[0.9842,1.0278]
GA+AA	522	1.0002[0.9983,1.0021]	1.0012[0.9934,1.0090]	1.0027[0.9938,1.0118]	1.0051[0.9921,1.0183]
rs3170633 G/A					
GG	366	Reference	Reference	Reference	Reference
GA	347	0.9994[0.9973,1.0016]	0.9975[0.9888,1.0062]	1.0016[0.9916,1.0117]	1.0001[0.9856,1.0147]
AA	171	0.9995[0.9970,1.0021]	0.9988[0.9884,1.0093]	1.0052[0.9933,1.0173]	1.0027[0.9851,1.0206]
GA+AA	518	0.9995[0.9975,1.0015]	0.9979[0.9900,1.0058]	1.0028[0.9937,1.0120]	1.0009[0.9876,1.0144]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01.

Table J7: Effect modification of SOD3 Genotypes on the association between 8-isoprostane and acute pollutant exposure (24hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9997[0.9980,1.0013]	0.9987[0.9924,1.0051]	0.9991[0.9919,1.0063]	0.9994[0.9887,1.0103]
TT	9	1.0115#[0.9980,1.0251]	1.0286#[0.9978,1.0604]	1.0343[0.9605,1.1139]	1.0629[0.9504,1.1888]
GT+TT	154	0.9998[0.9982,1.0015]	0.9998[0.9935,1.0061]	0.9995[0.9924,1.0067]	1.0002[0.9895,1.0110]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	1.0004[0.9991,1.0018]	1.0009[0.9954,1.0064]	1.0021[0.9958,1.0084]	1.0025[0.9935,1.0115]
TT	220	1.0009[0.9994,1.0023]	1.0016[0.9955,1.0078]	1.002[0.9951,1.0090]	1.003[0.9932,1.0128]
CT+TT	638	1.0006[0.9994,1.0019]	1.0012[0.9962,1.0063]	1.002[0.9964,1.0077]	1.0027[0.9947,1.0108]
rs13306703 C/T					
CC	622	Reference	Reference	Reference	Reference
CT	267	0.9999[0.9985,1.0012]	0.9999[0.9946,1.0052]	1.0017[0.9958,1.0076]	1.0021[0.9936,1.0107]
TT	40	0.9982[0.9952,1.0012]	0.9965[0.9851,1.0081]	0.9875#[0.9729,1.0023]	0.9839[0.9642,1.0041]
CT+TT	307	0.9996[0.9983,1.0009]	0.9993[0.9943,1.0044]	1.0001[0.9944,1.0058]	0.9998[0.9917,1.0080]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9998[0.9982,1.0014]	0.999[0.9927,1.0054]	0.9992[0.9920,1.0064]	1.0002[0.9895,1.0109]
TT	9	1.0116#[0.9981,1.0252]	1.0288#[0.9980,1.0605]	1.0343[0.9606,1.1138]	1.0629[0.9506,1.1885]
GT+TT	156	1[0.9983,1.0016]	1.0001[0.9938,1.0063]	0.9996[0.9925,1.0068]	1.001[0.9903,1.0117]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. *p<0.05. **p<0.01.

Table J8: Effect modification of SOD3 Genotypes on association between 8-oxodG and acute pollutant exposure (24hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	734	Reference	Reference	Reference	Reference
GT	138	1.0009[0.9981,1.0036]	1.0046[0.9939,1.0154]	1.0032[0.9912,1.0153]	1.0096[0.9911,1.0283]
TT	9	1.0001[0.9784,1.0223]	0.9948[0.9464,1.0457]	0.911[0.8076,1.0275]	0.8592[0.7169,1.0297]
GT+TT	147	1.0009[0.9981,1.0036]	1.0043[0.9938,1.0149]	1.0023[0.9904,1.0144]	1.008[0.9897,1.0266]
rs699473 C/T					
CC	276	Reference	Reference	Reference	Reference
CT	398	0.9982[0.9959,1.0005]	0.9939[0.9848,1.0030]	0.9914[0.9810,1.0018]	0.9872[0.9721,1.0025]
TT	209	0.9975*[0.9951,1.0000]	0.9884*[0.9783,0.9986]	0.9865*[0.9752,0.9979]	0.9812*[0.9651,0.9974]
CT+TT	607	0.9979*[0.9959,1.0000]	0.9918#[0.9835,1.0002]	0.9895*[0.9801,0.9989]	0.9848*[0.9713,0.9985]
rs13306703 C/T					
CC	587	Reference	Reference	Reference	Reference
CT	255	1.0018[0.9995,1.0041]	1.0084#[0.9996,1.0173]	1.0053[0.9954,1.0153]	1.0118[0.9972,1.0266]
TT	40	0.9974[0.9926,1.0023]	0.9853[0.9669,1.0041]	0.996[0.9721,1.0206]	0.9905[0.9584,1.0236]
CT+TT	295	1.0011[0.9990,1.0033]	1.0049[0.9966,1.0134]	1.0042[0.9947,1.0138]	1.0088[0.9949,1.0229]
rs8192288 G/T					
GG	737	Reference	Reference	Reference	Reference
GT	139	1.001[0.9983,1.0037]	1.005[0.9943,1.0157]	1.0033[0.9914,1.0154]	1.0101[0.9918,1.0288]
TT	9	1.0003[0.9786,1.0224]	0.9951[0.9467,1.0460]	0.9115[0.8081,1.0281]	0.8606[0.7182,1.0312]
GT+TT	148	1.001[0.9982,1.0037]	1.0046[0.9942,1.0152]	1.0025[0.9906,1.0145]	1.0086[0.9904,1.0271]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<o.1, *p<0.05, **p<0.01.

Table J9: Effect modification of Nrf2 and GCLM genotypes on the association between 8-isoprostane and sub-chronic pollutant exposure (7 day averages)

Genotype ₸	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	394	Reference	Reference	Reference	Reference
GC	382	1.0013[0.9979,1.0046]	1.0019[0.9928,1.0112]	1.0087[0.9979,1.0197]	1.0119[0.9959,1.0282]
CC	155	0.9976[0.9936,1.0018]	0.9941[0.9824,1.0060]	1.0056[0.9931,1.0183]	1.0095[0.9911,1.0281]
GC+CC	537	1.0001[0.9970,1.0031]	0.9996[0.9912,1.0080]	1.0077[0.9976,1.0179]	1.0111[0.9963,1.0261]
rs7557529 T/C					
TT	358	Reference	Reference	Reference	Reference
TC	426	1.0049**[1.0015,1.0082]	₸1.0151**[1.0058,1.0246]	1.0037[0.9938,1.0137]	1.0059[0.9913,1.0208]
CC	145	1.0018[0.9978,1.0058]	1.007[0.9955,1.0186]	0.9995[0.9852,1.0140]	0.9973[0.9763,1.0187]
TC+CC	571	1.0038*[1.0008,1.0068]	1.0125**[1.0040,1.0210]	1.0026[0.9935,1.0118]	1.0038[0.9902,1.0175]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	1[0.9962,1.0038]	1.0041[0.9934,1.0150]	1.0068[0.9960,1.0177]	1.0085[0.9923,1.0250]
GG	10	1.001[0.9852,1.0171]	0.9978[0.9552,1.0423]	0.9947[0.9655,1.0248]	0.9921[0.9496,1.0366]
AG+GG	210	1.0001[0.9964,1.0038]	1.0039[0.9934,1.0145]	1.0057[0.9953,1.0162]	1.0069[0.9913,1.0227]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	403	1.0021[0.9988,1.0054]	1.0046[0.9954,1.0139]	1.002[0.9924,1.0117]	1.0005[0.9862,1.0150]
AA	141	1.0006[0.9960,1.0052]	1.0019[0.9898,1.0142]	0.9846[0.9692,1.0002]	0.9824[0.9598,1.0055]
GA+AA	544	1.0017[0.9987,1.0048]	1.0038[0.9953,1.0124]	0.9987[0.9896,1.0079]	0.9971[0.9835,1.0108]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360	0.9987[0.9953,1.0022]	0.9963[0.9866,1.0060]	0.9863**[0.9767,0.9959]	0.9774**[0.9632,0.9917]
AA	184	1.0022[0.9982,1.0061]	1.0052[0.9942,1.0163]	0.9866#[0.9729,1.0005]	0.9800*[0.9604,1.0000]
GA+AA	544	1.0001[0.9970,1.0032]	0.9999[0.9911,1.0088]	0.9865**[0.9775,0.9955]	₸0.9783**[0.9652,0.9917]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01. ₸ Significant after the multiple testing procedures has been applied.

Table J10: Effect modification of Nrf2 and GCLM genotypes on the association between 8-oxodG and sub-chronic pollutant exposure (7 day averages)

Genotype ₸	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	374	Reference	Reference	Reference	Reference
GC	366	1.005#[0.9996,1.0105]	1.0184*[1.0031,1.0340]	1.0302**[1.0114,1.0493]	1.0318*[1.0041,1.0604]
CC	145	1.0029[0.9956,1.0101]	1.0105[0.9898,1.0315]	1.0212[0.9974,1.0456]	1.022[0.9873,1.0578]
GC+CC	511	1.0044#[0.9994,1.0095]	1.0163*[1.0022,1.0306]	₸1.0279**[1.0100,1.0460]	1.0290*[1.0029,1.0559]
rs7557529 T/C					
TT	340	Reference	Reference	Reference	Reference
TC	406	1.0041[0.9985,1.0097]	1.0169*[1.0013,1.0327]	1.0088[0.9913,1.0266]	1.0058[0.9800,1.0324]
CC	138	0.9969[0.9902,1.0036]	0.9909[0.9721,1.0101]	0.9787#[0.9552,1.0028]	0.9746[0.9403,1.0102]
TC+CC	544	1.0017[0.9966,1.0068]	1.0089[0.9947,1.0233]	1.0007[0.9846,1.0171]	0.9982[0.9743,1.0226]
rs2001350 A/G					
AA	684	Reference	Reference	Reference	Reference
AG	191	1.0006[0.9943,1.0069]	1.0133[0.9955,1.0315]	1.0014[0.9829,1.0201]	0.9989[0.9713,1.0273]
GG	10	0.9926[0.9672,1.0187]	0.9797[0.9126,1.0518]	0.9672[0.9209,1.0158]	0.9571[0.8906,1.0286]
AG+GG	201	0.9998[0.9937,1.0059]	1.0105[0.9931,1.0282]	0.9976[0.9799,1.0157]	0.9939[0.9674,1.0211]
			GCLM		
rs2301022 G/A					
GG	360	Reference	Reference	Reference	Reference
GA	385	1.0005[0.9951,1.0060]	0.9996[0.9843,1.0152]	0.9981[0.9802,1.0163]	1.0037[0.9770,1.0310]
AA	137	1.0005[0.9929,1.0081]	1.0045[0.9843,1.0250]	1.0061[0.9799,1.0331]	1.0037[0.9652,1.0437]
GA+AA	522	1.0006[0.9955,1.0057]	1.0012[0.9870,1.0156]	0.9999[0.9832,1.0168]	1.0033[0.9784,1.0289]
rs3170633 G/A					
GG	366	Reference	Reference	Reference	Reference
GA	347	0.9967[0.9910,1.0026]	0.9898[0.9736,1.0062]	0.9935[0.9763,1.0111]	0.9901[0.9645,1.0163]
AA	171	1.002[0.9955,1.0086]	0.9989[0.9808,1.0173]	1.009[0.9843,1.0343]	1.0142[0.9783,1.0514]
GA+AA	518	0.9989[0.9937,1.0041]	0.9937[0.9791,1.0084]	0.9976[0.9814,1.0140]	0.9967[0.9727,1.0212]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05,**p<0.01. ₸ Significant after the multiple testing procedures has been applied.

Table J11: Effect modification of SOD3 genotypes on association between 8-isoprostane and sub-chronic pollutant exposure (7 day averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9996[0.9956,1.0036]	0.9993[0.9879,1.0108]	0.9982[0.9857,1.0110]	0.9978[0.9784,1.0176]
TT	9	1.0135[0.9949,1.0324]	1.0485[0.9922,1.1080]	1.0096[0.9645,1.0569]	1.0225[0.9504,1.1000]
GT+TT	154	1.0001[0.9961,1.0040]	1.0009[0.9896,1.0123]	0.9988[0.9865,1.0112]	0.999[0.9801,1.0183]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	0.9982[0.9946,1.0017]	0.9928[0.9830,1.0026]	0.9859*[0.9749,0.9970]	0.9757**[0.9596,0.9920]
TT	220	0.9988[0.9946,1.0029]	0.9908[0.9795,1.0023]	0.9914[0.9792,1.0039]	0.9842#[0.9660,1.0027]
CT+TT	638	0.9985[0.9952,1.0017]	0.9923[0.9833,1.0015]	0.9880*[0.9777,0.9984]	0.9790**[0.9638,0.9943]
rs13306703 C/T					
CC	622	Reference	Reference	Reference	Reference
CT	267	1.0002[0.9968,1.0036]	0.9996[0.9904,1.0090]	1.0036[0.9937,1.0136]	1.0062[0.9913,1.0213]
TT	40	1.002[0.9943,1.0098]	1.011[0.9877,1.0349]	1.0129[0.9826,1.0440]	1.0241[0.9833,1.0665]
CT+TT	307	1.0004[0.9972,1.0036]	1.0007[0.9918,1.0097]	1.0042[0.9946,1.0140]	1.0078[0.9933,1.0224]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9999[0.9959,1.0039]	0.9999[0.9885,1.0114]	0.9999[0.9880,1.0119]	1.0011[0.9827,1.0199]
TT	9	1.0136[0.9951,1.0324]	1.0488#[0.9926,1.1082]	1.0099[0.9648,1.0570]	1.023[0.9510,1.1004]
GT+TT	156	1.0003[0.9964,1.0043]	1.0015[0.9903,1.0129]	1.0003[0.9887,1.0121]	1.0021[0.9841,1.0205]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01.

Table J12: Effect modification of SOD3 genotypes on association between 8-oxodG and sub-chronic pollutant exposure (7 day averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	734	Reference	Reference	Reference	Reference
GT	138	0.9982[0.9916,1.0048]	0.9987[0.9800,1.0178]	0.9915[0.9704,1.0130]	0.9798[0.9477,1.0130]
TT	9	1.0222[0.9919,1.0533]	1.0392[0.9496,1.1373]	1.0073[0.9349,1.0854]	1.0447[0.9274,1.1768]
GT+TT	147	0.9991[0.9927,1.0057]	1.0001[0.9816,1.0189]	0.9926[0.9721,1.0135]	0.984[0.9527,1.0164]
rs699473 C/T					
CC	276	Reference	Reference	Reference	Reference
CT	398	0.9985[0.9927,1.0044]	0.9985[0.9822,1.0150]	0.9992[0.9795,1.0192]	0.9936[0.9646,1.0235]
TT	209	0.9936#[0.9868,1.0005]	0.9800*[0.9616,0.9989]	0.9849[0.9639,1.0063]	0.9745[0.9434,1.0066]
CT+TT	607	0.9971[0.9916,1.0025]	0.9924[0.9774,1.0078]	0.9936[0.9755,1.0120]	0.9867[0.9599,1.0142]
rs13306703 C/T					
CC	587	Reference	Reference	Reference	Reference
CT	255	1.0033[0.9978,1.0089]	1.01[0.9946,1.0256]	0.9961[0.9790,1.0135]	0.9993[0.9736,1.0257]
TT	40	0.9837*[0.9714,0.9962]	0.9571*[0.9214,0.9941]	0.9766[0.9293,1.0264]	0.9531[0.8917,1.0187]
CT+TT	295	1.0009[0.9956,1.0062]	1.0043[0.9895,1.0193]	0.9946[0.9779,1.0115]	0.9947[0.9700,1.0201]
rs8192288 G/T					
GG	737	Reference	Reference	Reference	Reference
GT	139	0.9982[0.9917,1.0048]	0.9988[0.9801,1.0178]	0.9903[0.9696,1.0115]	0.9786[0.9466,1.0115]
TT	9	1.0223[0.9921,1.0534]	1.0397[0.9501,1.1377]	1.0075[0.9350,1.0855]	1.045[0.9278,1.1771]
GT+TT	148	0.9992[0.9927,1.0057]	1.0002[0.9817,1.0189]	0.9915[0.9713,1.0121]	0.9828[0.9516,1.0149]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05.

Table J13: Effects modification of Nrf2 and GCLM genotypes on the association between 8-isoprostane and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	394	Reference	Reference	Reference	Reference
GC	382	1.0006[0.9966,1.0046]	1.0018[0.9912,1.0125]	1.0069[0.9686,1.0469]	1.029[0.9623,1.1002]
CC	155	0.9987[0.9945,1.0029]	0.9967[0.9853,1.0082]	0.9897[0.9452,1.0363]	0.9978[0.9170,1.0858]
GC+CC	537	0.9997[0.9962,1.0033]	0.9997[0.9903,1.0092]	1.0009[0.9661,1.0370]	1.0193[0.9587,1.0837]
rs7557529 T/C					
TT	358	Reference	Reference	Reference	Reference
TC	426	1.0015[0.9973,1.0057]	1.0043[0.9934,1.0154]	1.0039[0.9659,1.0434]	0.9675[0.9049,1.0344]
CC	145	1.001[0.9967,1.0054]	1.0025[0.9906,1.0146]	1.0054[0.9565,1.0568]	0.9755[0.8939,1.0646]
TC+CC	571	1.0013[0.9978,1.0047]	1.0035[0.9944,1.0128]	1.0044[0.9701,1.0398]	0.9699[0.9120,1.0315]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	0.9999[0.9945,1.0053]	0.9996[0.9856,1.0139]	0.9907[0.9423,1.0415]	0.9732[0.8975,1.0553]
GG	10	1.0001[0.9844,1.0161]	0.9999[0.9610,1.0404]	1.0032[0.8737,1.1520]	0.9832[0.7701,1.2552]
AG+GG	210	0.9999[0.9947,1.0050]	0.9996[0.9862,1.0131]	0.9918[0.9455,1.0404]	0.9737[0.9007,1.0527]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	403	1.0008[0.9973,1.0044]	1.0023[0.9928,1.0118]	1.0132[0.9772,1.0505]	1.0199[0.9556,1.0885]
AA	141	0.9957[0.9887,1.0028]	0.9888[0.9710,1.0070]	0.9424#[0.8852,1.0032]	0.9424[0.8558,1.0378]
GA+AA	544	1.0001[0.9967,1.0035]	1.0002[0.9912,1.0094]	1.001[0.9668,1.0365]	1.0022[0.9422,1.0660]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360	0.9996[0.9957,1.0035]	0.999[0.9885,1.0095]	1.0003[0.9622,1.0400]	1.0147[0.9476,1.0865]
AA	184	1.0015[0.9971,1.0058]	1.0038[0.9923,1.0155]	1.0199[0.9748,1.0670]	1.0158[0.9375,1.1005]
GA+AA	544	1.0004[0.9968,1.0039]	1.0011[0.9917,1.0105]	1.0075[0.9727,1.0435]	1.0154[0.9544,1.0803]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1.

Table J14: Effects modification of Nrf2 and GCLM genotypes on association between 8-oxodG and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	374	Reference	Reference	Reference	Reference
GC	366	1.0035[0.9968,1.0102]	1.0101[0.9924,1.0282]	1.0329[0.9676,1.1026]	1.0143[0.9058,1.1358]
CC	145	1.0014[0.9936,1.0094]	1.005[0.9842,1.0262]	1.0196[0.9420,1.1037]	0.9994[0.8658,1.1536]
GC+CC	511	1.0028[0.9967,1.0088]	1.0084[0.9923,1.0247]	1.0286[0.9690,1.0919]	1.0101[0.9108,1.1203]
rs7557529 T/C					
TT	340	Reference	Reference	Reference	Reference
TC	406	1.0027[0.9955,1.0100]	1.0079[0.9892,1.0268]	1.0201[0.9556,1.0890]	1.0137[0.9056,1.1348]
CC	138	1.002[0.9947,1.0095]	1.0049[0.9848,1.0253]	1.0418[0.9577,1.1332]	1.1124[0.9605,1.2884]
TC+CC	544	1.0024[0.9964,1.0084]	1.0066[0.9909,1.0226]	1.026[0.9674,1.0882]	1.0411[0.9382,1.1553]
rs2001350 A/G					
AA	684	Reference	Reference	Reference	Reference
AG	191	1.0009[0.9918,1.0100]	1.0031[0.9796,1.0272]	1.0096[0.9288,1.0975]	0.9769[0.8530,1.1188]
GG	10	1.0098[0.9841,1.0362]	1.026[0.9617,1.0946]	1.1235[0.8968,1.4075]	1.3075[0.8784,1.9461]
AG+GG	201	1.0022[0.9935,1.0109]	1.0065[0.9840,1.0296]	1.0234[0.9449,1.1083]	1.0086[0.8849,1.1495]
			GCLM		
rs2301022 G/A					
GG	360	Reference	Reference	Reference	Reference
GA	385	0.9977[0.9913,1.0042]	0.9939[0.9773,1.0108]	0.9778[0.9192,1.0402]	0.9765[0.8742,1.0907]
AA	137	1.0018[0.9898,1.0140]	1.0033[0.9727,1.0349]	0.9836[0.8851,1.0930]	0.9635[0.8193,1.1331]
GA+AA	522	0.9985[0.9925,1.0046]	0.9958[0.9800,1.0118]	0.9797[0.9237,1.0391]	0.974[0.8775,1.0811]
rs3170633 G/A					
GG	366	Reference	Reference	Reference	Reference
GA	347	1.0043[0.9971,1.0114]	1.0124[0.9937,1.0315]	1.0399[0.9730,1.1115]	1.0976[0.9775,1.2323]
AA	171	1.0014[0.9943,1.0086]	1.0034[0.9844,1.0228]	1.0056[0.9330,1.0840]	0.9975[0.8723,1.1406]
GA+AA	518	1.0029[0.9969,1.0090]	1.0082[0.9922,1.0244]	1.0263[0.9672,1.0891]	1.0593[0.9542,1.1759]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.

Table J15: Effects modification of SOD3 genotypes on association between 8-isoprostane and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9991[0.9943,1.0038]	0.9973[0.9847,1.0101]	0.9935[0.9483,1.0410]	0.9963[0.9158,1.0839]
TT	9	1.0351#[0.9937,1.0783]	1.0799[0.9778,1.1927]	1.3887[0.8088,2.3844]	1.1988[0.6625,2.1692]
GT+TT	154	0.9996[0.9949,1.0043]	0.9987[0.9861,1.0114]	0.9972[0.9520,1.0445]	1.0033[0.9233,1.0904]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	0.9987[0.9942,1.0033]	0.9968[0.9849,1.0088]	0.9961[0.9536,1.0405]	0.9863[0.9165,1.0615]
TT	220	0.9993[0.9943,1.0043]	0.9979[0.9848,1.0111]	1.0009[0.9554,1.0485]	0.9849[0.9067,1.0699]
CT+TT	638	0.9989[0.9946,1.0032]	0.9971[0.9858,1.0085]	0.998[0.9586,1.0390]	0.9854[0.9204,1.0551]
rs13306703 C/T					
CC	622	Reference	Reference	Reference	Reference
CT	267	1.0003[0.9963,1.0044]	1.001[0.9901,1.0119]	1.0096[0.9707,1.0500]	1.0288[0.9602,1.1022]
TT	40	1.0018[0.9841,1.0199]	1.0075[0.9613,1.0559]	1.0587[0.8748,1.2813]	0.9978[0.7965,1.2500]
CT+TT	307	1.0003[0.9963,1.0043]	1.001[0.9903,1.0117]	1.0096[0.9713,1.0495]	1.0237[0.9573,1.0946]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9989[0.9942,1.0036]	0.997[0.9844,1.0097]	0.9919[0.9470,1.0389]	0.9902[0.9108,1.0764]
TT	9	1.0353#[0.9940,1.0784]	1.0806[0.9786,1.1932]	1.3917[0.8114,2.3870]	1.1966[0.6620,2.1629]
GT+TT	156	0.9994[0.9947,1.0041]	0.9984[0.9858,1.0111]	0.9956[0.9508,1.0425]	0.9972[0.9182,1.0830]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1.

Table J16: Effects modification of SOD3 genotypes on association between 8-oxodG and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
rs8192287 G/T					
GG	734	Reference	Reference	Reference	Reference
GT	138	1.0045[0.9963,1.0127]	1.0114[0.9897,1.0337]	1.0446[0.9620,1.1344]	1.0983[0.9477,1.2727]
TT	9	0.989[0.9251,1.0573]	0.9802[0.8333,1.1531]	1.1524[0.4766,2.7860]	0.9214[0.3507,2.4207]
GT+TT	147	1.0042[0.9961,1.0124]	1.0108[0.9893,1.0329]	1.0445[0.9624,1.1335]	1.0923[0.9445,1.2632]
rs699473 C/T					
CC	276	Reference	Reference	Reference	Reference
CT	398	1.002[0.9943,1.0099]	1.0044[0.9842,1.0250]	1.0177[0.9452,1.0959]	1.0306[0.9098,1.1674]
TT	209	0.9993[0.9910,1.0076]	0.9971[0.9757,1.0191]	0.9865[0.9132,1.0657]	1.0125[0.8810,1.1636]
CT+TT	607	1.0009[0.9936,1.0082]	1.0013[0.9825,1.0205]	1.0037[0.9381,1.0739]	1.0236[0.9119,1.1489]
rs13306703 C/T					
CC	587	Reference	Reference	Reference	Reference
CT	255	0.9994[0.9925,1.0064]	0.9977[0.9795,1.0163]	0.9845[0.9202,1.0532]	0.9679[0.8600,1.0893]
TT	40	0.9947[0.9659,1.0242]	0.9902[0.9169,1.0694]	0.9685[0.7089,1.3230]	1.0122[0.7006,1.4625]
CT+TT	295	0.9991[0.9923,1.0060]	0.9971[0.9792,1.0153]	0.9825[0.9194,1.0499]	0.9682[0.8634,1.0857]
rs8192288 G/T					
GG	737	Reference	Reference	Reference	Reference
GT	139	1.0046[0.9964,1.0128]	1.0118[0.9901,1.0340]	1.0472[0.9648,1.1367]	1.1051[0.9545,1.2793]
TT	9	0.9887[0.9248,1.0569]	0.9793[0.8327,1.1518]	1.1445[0.4738,2.7646]	0.9162[0.3490,2.4049]
GT+TT	148	1.0043[0.9962,1.0125]	1.0112[0.9897,1.0332]	1.0471[0.9652,1.1358]	1.099[0.9512,1.2699]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school.

Table J17: Effect modification of Nrf2 and GCLM Genotypes on the association between Cu and acute pollutant exposure (24 hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	387	Reference	Reference	Reference	Reference
GC	377	1.0000[0.9999,1.0001]	1.0002[0.9998,1.0007]	1.0001[0.9997,1.0005]	1.0002[0.9996,1.0008]
CC	148	1.0001[1.0000,1.0003]	1.0005#[1.0000,1.0011]	1.0002[0.9996,1.0009]	1.0006[0.9996,1.0015]
GC+CC	525	1.0001[1.0000,1.0002]	1.0001[0.9997,1.0005]	1.0001[0.9997,1.0005]	1.0003[0.9997,1.0009]
rs7557529 T/C					
TT	351	Reference	Reference	Reference	Reference
TC	418	1.0000[0.9999,1.0001]	0.9999[0.9995,1.0003]	0.9999[0.9994,1.0004]	0.9998[0.9992,1.0005]
CC	141	0.9999[0.9998,1.0001]	0.9996[0.9990,1.0001]	0.9995[0.9989,1.0002]	0.9995[0.9987,1.0004]
TC+CC	559	1.0000[0.9999,1.0001]	0.9998[0.9994,1.0002]	0.9999[0.9994,1.0003]	0.9998[0.9991,1.0004]
rs2001350 A/G					
AA	707				
AG	195	1.0000[0.9998,1.0001]	0.9997[0.9993,1.0002]	0.9997[0.9992,1.0002]	0.9995[0.9989,1.0002]
GG	10	1.0000[0.9994,1.0006]	1.0004[0.9983,1.0025]	0.9999[0.9971,1.0026]	0.9988[0.9944,1.0032]
AG+GG	205	1.0000[0.9999,1.0001]	0.9998[0.9993,1.0002]	0.9997[0.9992,1.0002]	0.9995[0.9988,1.0002]
			GCLM		
rs2301022 G/A					
GG	375	Reference	Reference	Reference	Reference
GA	394	1.0000[0.9999,1.0001]	1.0001[0.9997,1.0005]	1.0000[0.9996,1.0005]	1.0000[0.9994,1.0006]
AA	140	1.0001[0.9999,1.0003]	1.0006#[1.0000,1.0012]	1.0005[0.9998,1.0012]	1.0003[0.9993,1.0013]
GA+AA	534	1.0000[0.9999,1.0001]	1.0002[0.9998,1.0006]	1.0001[0.9997,1.0006]	1.0001[0.9995,1.0007]
rs3170633 G/A					
GG	380	Reference	Reference	Reference	Reference
GA	352	1.0000[0.9999,1.0001]	0.9998[0.9994,1.0002]	0.9997[0.9993,1.0002]	0.9996[0.9989,1.0003]
AA	179	0.9999[0.9998,1.0001]	0.9997[0.9992,1.0002]	0.9996[0.9991,1.0002]	0.9993#[0.9986,1.0001]
GA+AA	531	1.0000[0.9999,1.0000]		0.9997[0.9993,1.0001]	0.9994#[0.9988,1.0001]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1.
Notes: Blank indicates non-converged data.

Table J18: Effect modification of Nrf2 and GCLM Genotypes on the association between Ni and acute pollutant exposure (24hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	393	Reference	Reference	Reference	Reference
GC	383	1.0001[0.9997,1.0004]	1.0004[0.9987,1.0020]	1.0003[0.9985,1.0020]	1.0005[0.9982,1.0029]
CC	155	1.0004[0.9998,1.0010]	1.0016[0.9993,1.0039]	1.0014[0.9989,1.0040]	1.0031[0.9993,1.0069]
GC+CC	538	1.0001[0.9998,1.0005]	1.0006[0.9992,1.0021]	1.0006[0.9989,1.0022]	1.0010[0.9988,1.0033]
rs7557529 T/C					
TT	359	Reference	Reference	Reference	Reference
TC	425	0.9999[0.9995,1.0003]	0.9996[0.9979,1.0013]	0.9986[0.9966,1.0005]	0.9983[0.9956,1.0009]
CC	145	0.9997[0.9992,1.0002]	0.999[0.9969,1.0011]	0.9979[0.9956,1.0003]	0.998[0.9948,1.0012]
TC+CC	570	0.9998[0.9994,1.0002]	0.9994[0.9979,1.0010]	0.9990[0.9972,1.0008]	0.9974#[0.9948,1.0001]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	1.0000[0.9996,1.0004]	1.0000[0.9983,1.0017]	0.9995[0.9977,1.0014]	0.9994[0.9969,1.0019]
GG	10	0.9999[0.9979,1.0020]	1.0008[0.9935,1.0082]	0.9975[0.9878,1.0072]	0.9945[0.9798,1.0095]
AG+GG	210	1.0000[0.9996,1.0004]	0.9997[0.9979,1.0015]	0.9994[0.9975,1.0013]	0.9991[0.9965,1.0018]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	402	0.9999[0.9996,1.0003]	0.9996[0.9979,1.0013]	0.9996[0.9978,1.0014]	0.9998[0.9974,1.0021]
AA	142	1.0000[0.9993,1.0006]	1.0004[0.9980,1.0029]	1.0003[0.9974,1.0032]	0.9996[0.9955,1.0038]
GA+AA	544	0.9999[0.9996,1.0003]	0.9998[0.9982,1.0013]	0.9997[0.9980,1.0014]	0.9997[0.9975,1.0020]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360	0.9999[0.9995,1.0003]	0.9995[0.9978,1.0011]	0.9991[0.9972,1.0009]	0.9982[0.9957,1.0008]
AA	184	0.9999[0.9995,1.0004]	0.999[0.9971,1.0010]	0.9985[0.9964,1.0006]	0.9979[0.9949,1.0009]
GA+AA	544	0.9999[0.9995,1.0003]	0.9993[0.9978,1.0008]	0.9989[0.9972,1.0006]	0.9981[0.9957,1.0005]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1.

Table J19: Effect modification of SOD3 Genotypes on the association between Cu and acute pollutant exposure (24hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	754	Reference	Reference	Reference	Reference
GT	145	0.9999[0.9998,1.0001]	0.9999[0.9994,1.0005]	0.9996[0.9990,1.0001]	0.9994[0.9985,1.0002]
TT	9	1.0002[0.9991,1.0012]	1.0007[0.9981,1.0032]	1.0041[0.9986,1.0096]	1.0039[0.9952,1.0126]
GT+TT	154	1.0000[0.9998,1.0001]	0.9999[0.9994,1.0004]	0.9996[0.9991,1.0002]	0.9994[0.9986,1.0003]
rs699473 C/T					
CC	754	Reference	Reference	Reference	Reference
CT	145	1.0001[1.0000,1.0002]	1.0003[0.9998,1.0007]		1.0005[0.9998,1.0012]
TT	9	1.0001*[1.0000,1.0002]	1.0003[0.9998,1.0008]		1.0006[0.9999,1.0013]
CT+TT	154	1.0001[1.0000,1.0002]	1.0003[0.9999,1.0007]	1.0003[0.9999,1.0008]	1.0006#[0.9999,1.0012]
rs13306703 C/T					
CC	606	Reference	Reference	Reference	Reference
CT	264	0.9999[0.9998,1.0000]	0.9999[0.9994,1.0003]	0.9996#[0.9991,1.0001]	0.9995[0.9988,1.0001]
TT	39	1.0000[0.9997,1.0002]	0.9998[0.9989,1.0008]	0.9996[0.9984,1.0009]	0.9999[0.9984,1.0015]
CT+TT	303	0.9999[0.9998,1.0000]	0.9999[0.9995,1.0003]	0.9996#[0.9991,1.0001]	0.9994#[0.9988,1.0001]
rs8192288 G/T					
GG	757	Reference	Reference	Reference	Reference
GT	146	1.0000[0.9998,1.0001]	0.9999[0.9994,1.0004]	0.9995[0.9990,1.0001]	0.9995[0.9986,1.0003]
TT	9	1.0002[0.9991,1.0012]	1.0005[0.9981,1.0030]	1.0041[0.9986,1.0096]	1.0039[0.9952,1.0126]
GT+TT	155	1.0000[0.9998,1.0001]	1.0000[0.9995,1.0006]	0.9996[0.9990,1.0001]	0.9995[0.9987,1.0004]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05. Notes: Blank indicates non-converged data.

Table J20: Effect modification of SOD3 Genotypes on the association between Ni and acute pollutant exposure (24hour)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9999[0.9994,1.0004]	0.9993[0.9972,1.0013]	0.9982[0.9960,1.0004]	0.9978[0.9948,1.0009]
TT	9	0.9981[0.9943,1.0020]	0.996[0.9873,1.0048]	0.9935[0.9727,1.0146]	0.9945[0.9636,1.0264]
GT+TT	154	0.9997[0.9994,1.0003]	0.9996[0.9975,1.0017]	0.9987[0.9964,1.0010]	0.9980[0.9949,1.0010]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	1.0000[0.9996,1.0004]	1.0002[0.9985,1.0020]	1.0009[0.9990,1.0028]	1.0015[0.9987,1.0042]
TT	220	1.0002[0.9998,1.0006]	1.0007[0.9988,1.0026]	1.0015[0.9994,1.0035]	1.0026[0.9996,1.0055]
CT+TT	638	1.0001[0.9998,1.0005]	1.0005[0.9988,1.0021]	1.0011[0.9994,1.0028]	1.0014[0.9991,1.0038]
rs13306703 C/T					
CC	621	Reference	Reference	Reference	Reference
CT	267	0.9999[0.9995,1.0003]	0.9995[0.9978,1.0012]	0.9983[0.9965,1.0001]	0.9982[0.9957,1.0006]
TT	40	1.0001[0.9993,1.0010]	1.0008[0.9975,1.0041]	1.0001[0.9960,1.0042]	1.0011[0.9955,1.0067]
CT+TT	307	0.9999[0.9995,1.0003]	0.9998[0.9981,1.0015]	0.9986[0.9969,1.0003]	0.9986[0.9963,1.0010]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9999[0.9994,1.0005]	0.9993[0.9972,1.0013]	0.9982[0.9960,1.0003]	0.9977[0.9947,1.0008]
TT	9	0.9994[0.9953,1.0035]	0.9962[0.9875,1.0050]	0.9933[0.9725,1.0145]	0.9945[0.9636,1.0264]
GT+TT	156	0.9999[0.9994,1.0003]	0.9995[0.9974,1.0016]	0.9982[0.9960,1.0004]	0.9979[0.9948,1.0009]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. Notes: Blank indicates non-converged data

Table J21: Effect modification of Nrf2 and GCLM genotypes on the association between Cu and sub-acute pollutant exposure (7 days averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	387	Reference	Reference	Reference	Reference
GC	377	1.0001[0.9998,1.0003]	0.9999[0.9992,1.0007]	1.0005[0.9996,1.0014]	1.0014*[1.0000,1.0027]
CC	148	1.0002[0.9999,1.0006]	1.0006[0.9996,1.0016]	1.0006[0.9994,1.0018]	1.0018*[1.0000,1.0035]
GC+CC	525	1.0001[0.9998,1.0003]	1.0001[0.9994,1.0008]	1.0005[0.9997,1.0014]	1.0008[0.9996,1.0020]
rs7557529 T/C					
TT	351	Reference	Reference	Reference	Reference
TC	418	0.9998[0.9995,1.0001]	0.9996[0.9988,1.0004]	0.9995[0.9987,1.0004]	0.9987[0.9975,1.0000]
CC	141	1.0000[0.9996,1.0003]	0.9999[0.9989,1.0009]	0.9995[0.9983,1.0007]	0.9994[0.9976,1.0012]
TC+CC	559	0.9999[0.9996,1.0001]	0.9995[0.9989,1.0002]	0.9996[0.9988,1.0004]	
rs2001350 A/G					
AA	707	Reference	Reference	Reference	Reference
AG	195	0.9999[0.9996,1.0003]	0.9999[0.9990,1.0007]	0.9998[0.9989,1.0008]	0.9994[0.9980,1.0007]
GG	10	0.9997[0.9984,1.0010]	1.0002[0.9969,1.0035]	1.0004[0.9979,1.0030]	1.0005[0.9969,1.0042]
AG+GG	205	0.9999[0.9996,1.0002]	0.9999[0.9991,1.0008]	1.0001[0.9992,1.0010]	0.9995[0.9982,1.0008]
			GCLM		
rs2301022 G/A					
GG	375	Reference	Reference	Reference	Reference
GA	394	1.0000[0.9997,1.0002]	1.0000[0.9992,1.0007]	0.9994[0.9985,1.0003]	0.9994[0.9981,1.0007]
AA	140	0.9999[0.9996,1.0003]	1.0003[0.9993,1.0012]	0.9988#[0.9975,1.0001]	0.9977*[0.9958,0.9996]
GA+AA	534	1.0000[0.9998,1.0003]	1.0001[0.9994,1.0009]	0.9993#[0.9984,1.0001]	0.9992[0.9980,1.0004]
rs3170633 G/A					
GG	380	Reference	Reference	Reference	Reference
GA	352	0.9999[0.9996,1.0001]	0.9996[0.9988,1.0004]	0.9994[0.9986,1.0003]	0.999[0.9977,1.0003]
AA	179	1.0001[0.9998,1.0004]	1.0004[0.9995,1.0013]	0.9993[0.9980,1.0005]	0.9986[0.9968,1.0003]
GA+AA	531	0.9999[0.9997,1.0002]	0.9998[0.9991,1.0005]	0.9994[0.9986,1.0002]	0.9989#[0.9977,1.0001]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.5. Notes: Blank cells indicates non-converged data.

Table J22: Effect modification of Nrf2 and GCLM genotypes on the association between Ni and sub-acute pollutant exposure (7 days averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	393	Reference	Reference	Reference	Reference
GC	383	0.9998[0.9987,1.0009]	1.0002[0.9972,1.0031]	1.002[0.9986,1.0055]	1.0029[0.9981,1.0077]
CC	155	1.0006[0.9992,1.0021]	1.0017[0.9976,1.0057]	1.0019[0.9973,1.0065]	1.0036[0.9974,1.0099]
GC+CC	538	1.0000[0.9993,1.0012]	1.0002[0.9973,1.0031]	1.0020[0.9987,1.0053]	1.0032[0.9986,1.0077]
rs7557529 T/C					
TT	359	Reference	Reference	Reference	Reference
TC	425	0.9990[0.9979,1.0001]	0.9976[0.9945,1.0007]	0.9979[0.9946,1.0012]	0.9956[0.9909,1.0002]
CC	145	0.9995[0.9983,1.0008]	0.9986[0.9949,1.0024]	0.9982[0.9937,1.0027]	0.9979[0.9916,1.0042]
TC+CC	570	0.9992[0.9982,1.0002]	0.9979[0.9950,1.0007]	0.9980[0.9949,1.0010]	0.9961#[0.9918,1.0004]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	0.9991[0.9979,1.0003]	0.9978[0.9944,1.0013]	0.9962*[0.9928,0.9995]	
GG	10	0.9997[0.9954,1.0040]	0.9982[0.9866,1.0099]	0.9985[0.9898,1.0073]	
AG+GG	210	0.9990[0.9978,1.0002]	0.9978[0.9945,1.0012]	0.9964*[0.9931,0.9997]	0.9956#[0.9909,1.0003]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	402	0.9998[0.9988,1.0009]	0.9997[0.9966,1.0028]	0.9987[0.9952,1.0022]	0.9988[0.9940,1.0037]
AA	142	1.0001[0.9987,1.0016]	1.0020[0.9981,1.0060]	0.9979[0.9926,1.0033]	0.9949[0.9878,1.0021]
GA+AA	544	0.9999[0.9989,1.0009]	1.0006[0.9977,1.0035]	0.9985[0.9953,1.0017]	0.9979[0.9934,1.0025]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360	1.0001[0.9990,1.0013]	1.0007[0.9975,1.0040]	0.9998[0.9964,1.0031]	0.9987[0.9939,1.0035]
AA	184	1.0011[0.9998,1.0024]	1.0024[0.9988,1.0061]	1.0009[0.9961,1.0058]	1.0009[0.9936,1.0082]
GA+AA	544	1.0004[0.9993,1.0014]	1.0013[0.9984,1.0043]	1.000[0.9969,1.0031]	

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.5. Notes: Blank cells indicates non-converged data.

Table J23: Effect modification of SOD3 genotypes on the association between Cu and sub-acute pollutant exposure (7 days averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	754	Reference	Reference	Reference	Reference
GT	145	0.9998[0.9995,1.0001]	0.9993[0.9984,1.0002]	0.9997[0.9987,1.0008]	0.9996[0.9980,1.0011]
TT	9	1.0001[0.9986,1.0016]	1.0002[0.9959,1.0046]	1.0017[0.9976,1.0058]	1.0024[0.9967,1.0082]
GT+TT	154	0.9998[0.9995,1.0001]	0.9995[0.9985,1.0004]	0.9998[0.9988,1.0008]	0.9998[0.9983,1.0013]
rs699473 C/T					
CC	754	Reference	Reference	Reference	Reference
CT	145	1.0001[0.9998,1.0004]	1.0003[0.9994,1.0011]	1.0002[0.9992,1.0012]	1.0004[0.9990,1.0018]
TT	9	1.0002[0.9999,1.0006]	1.0004[0.9995,1.0014]	1.0013*[1.0002,1.0023]	1.0019*[1.0004,1.0035]
CT+TT	154	1.0002[0.9999,1.0005]	1.0002[0.9995,1.0010]	1.0007[0.9997,1.0016]	1.0012#[0.9999,1.0025]
rs13306703 C/T					
CC	606	Reference	Reference	Reference	Reference
CT	264	0.9998[0.9995,1.0001]	0.9995[0.9987,1.0003]	0.9995[0.9986,1.0004]	0.9991[0.9978,1.0004]
TT	39	0.9998[0.9992,1.0004]	0.999[0.9970,1.0010]	0.9995[0.9969,1.0020]	0.9985[0.9953,1.0019]
CT+TT	303	0.9998#[0.9995,1.0000]	0.9995[0.9988,1.0003]	0.9995[0.9986,1.0003]	0.999[0.9977,1.0003]
rs8192288 G/T					
GG	757	Reference	Reference	Reference	Reference
GT	146	0.9998[0.9995,1.0002]	0.9994[0.9985,1.0003]	0.9993[0.9983,1.0003]	0.9992[0.9977,1.0008]
TT	9	1.0001[0.9986,1.0016]	1.0001[0.9957,1.0044]	1.0016[0.9975,1.0057]	1.0024[0.9967,1.0081]
GT+TT	155	0.9998[0.9995,1.0001]	0.9995[0.9986,1.0005]	0.9993[0.9983,1.0003]	0.9996[0.9980,1.0012]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05, **p<0.01.

Table J24: Effect modification of SOD3 genotypes on the association between Ni and sub-acute pollutant exposure (7 days averages)

Genotype	Subjects (n)	NOx Exp(β) (95% CI)	NO_2 Exp(β) (95% CI)	PM_{10} Exp(β) (95% CI)	$PM_{2.5}$ Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9994[0.9981,1.0006]		0.9985[0.9947,1.0023]	0.999[0.9932,1.0049]
TT	9	0.9960[0.9904,1.0016]		0.9992[0.9844,1.0142]	1.0013[0.9793,1.0237]
GT+TT	154	0.9993[0.9980,1.0007]	0.9972[0.9935,1.0009]	0.9985[0.9948,1.0022]	0.9992[0.9936,1.0049]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	1.0000[0.9988,1.0012]	0.9998[0.9966,1.0030]	1.0004[0.9965,1.0044]	1.001[0.9956,1.0063]
TT	220	1.0007[0.9993,1.0021]	1.0025[0.9988,1.0061]	1.0033[0.9991,1.0074]	1.0046[0.9989,1.0104]
CT+TT	638	1.0004[0.9993,1.0014]	1.0007[0.9978,1.0037]	1.0017[0.9981,1.0053]	1.0024[0.9974,1.0074]
rs13306703 C/T					
CC	621	Reference	Reference	Reference	Reference
CT	267	0.9998[0.9987,1.0010]	0.999[0.9960,1.0021]	0.9986[0.9952,1.0020]	0.9991[0.9944,1.0039]
TT	40	1.0001[0.9978,1.0025]	1.0002[0.9935,1.0069]	0.9987[0.9893,1.0082]	1.0014[0.9900,1.0129]
CT+TT	307		0.9993[0.9962,1.0024]	0.9986[0.9953,1.0019]	0.9994[0.9948,1.0039]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9993[0.9980,1.0006]	0.9973[0.9936,1.0011]	0.9984[0.9944,1.0024]	0.9989[0.9931,1.0047]
TT	9	0.996[0.9904,1.0015]	0.9897[0.9737,1.0060]	1.0020[0.9865,1.0177]	1.0012[0.9793,1.0237]
GT+TT	156	0.9991[0.9979,1.0004]	0.9971[0.9934,1.0008]		0.999[0.9934,1.0047]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. Notes: Blank cells indicate non-converged data.

Table J25: Effects modification of Nrf2 and GCLM genotypes on the association between Cu and annual mean pollutant attributions

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	387	Reference	Reference	Reference	Reference
GC	377	0.9998[0.9995,1.0001]	0.9994[0.9985,1.0003]	0.9988[0.9957,1.0019]	0.9959[0.9903,1.0015]
CC	148	1.0000[0.9996,1.0004]	0.9997[0.9987,1.0008]	0.9995[0.9958,1.0033]	0.9965[0.9893,1.0037]
GC+CC	525	0.9999[0.9996,1.0002]	0.9997[0.9990,1.0005]	0.9990[0.9962,1.0019]	0.9967[0.9918,1.0017]
rs7557529 T/C					
TT	351	Reference	Reference	Reference	Reference
TC	418	1.0002[0.9998,1.0005]	1.0002[0.9993,1.0011]	1.0012[0.9981,1.0043]	1.0009[0.9956,1.0063]
CC	141	1.0001[0.9998,1.0005]	1.0001[0.9991,1.0011]	1.0011[0.9970,1.0052]	1.0026[0.9955,1.0097]
TC+CC	559	1.0001[0.9998,1.0003]	1.0004[0.9997,1.0012]	1.0013[0.9984,1.0042]	1.0015[0.9965,1.0065]
rs2001350 A/G					
AA	707	Reference	Reference	Reference	Reference
AG	195	1.0003[0.9999,1.0008]	1.0008[0.9997,1.0020]	1.0023[0.9979,1.0066]	1.0026[0.9957,1.0095]
GG	10	0.9996[0.9984,1.0008]	0.9997[0.9968,1.0027]	0.9964[0.9857,1.0073]	0.9954[0.9757,1.0154]
AG+GG	205	1.0003[0.9999,1.0007]	1.0006[0.9994,1.0017]	1.0023[0.9983,1.0062]	1.0020[0.9953,1.0086]
			GCLM		
rs2301022 G/A					
GG	375	Reference	Reference	Reference	Reference
GA	394	1.0000[0.9997,1.0003]	1.0002[0.9993,1.0010]	1.0005[0.9976,1.0035]	1.0020[0.9965,1.0075]
AA	140	1.0005#[1.0000,1.0011]	1.0006[0.9991,1.0021]	1.0050*[1.0001,1.0100]	1.0063[0.9983,1.0145]
GA+AA	534	1.0001[0.9998,1.0004]	1.0003[0.9996,1.0011]	1.0012[0.9982,1.0041]	1.0040[0.9990,1.0090]
rs3170633 G/A					
GG	380	Reference	Reference	Reference	Reference
GA	352	1.0000[0.9997,1.0003]	1.0000[0.9991,1.0009]	1.0006[0.9973,1.0039]	1.0005[0.9951,1.0059]
AA	179	1.0001[0.9997,1.0004]	1.0004[0.9994,1.0013]	1.0012[0.9974,1.0050]	0.9996[0.9930,1.0062]
GA+AA	531	1.0001[0.9998,1.0004]	1.0000[0.9993,1.0008]	1.0008[0.9979,1.0038]	1.0002[0.9953,1.0051]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.5.

Table J26: Effects modification of Nrf2 and GCLM genotypes on the association between Ni and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
			Nrf2		
rs2364723 G/C					
GG	393	Reference	Reference	Reference	Reference
GC	383	0.9992[0.9977,1.0008]	0.9973[0.9932,1.0014]	0.9918[0.9781,1.0058]	0.9788[0.9567,1.0014]
CC	155	1.0003[0.9983,1.0022]	1.0011[0.9960,1.0063]	1.0030[0.9852,1.0213]	0.9926[0.9635,1.0226]
GC+CC	538	0.9996[0.9982,1.0009]		0.9953[0.9827,1.0081]	0.9829[0.9626,1.0036]
rs7557529 T/C					
TT	359	Reference	Reference	Reference	Reference
TC	425	0.9997[0.9980,1.0014]	0.9993[0.9951,1.0036]		
CC	145	1.0001[0.9982,1.0019]	1.0003[0.9955,1.0052]		
TC+CC	570	0.9998[0.9984,1.0013]	0.9997[0.9959,1.0035]	0.9987[0.9854,1.0122]	1.0008[0.9788,1.0232]
rs2001350 A/G					
AA	721	Reference	Reference	Reference	Reference
AG	200	1.0004[0.9983,1.0024]	1.0005[0.9954,1.0057]	1.0019[0.9842,1.0199]	1.0037[0.9760,1.0322]
GG	10	0.9969[0.9930,1.0009]	0.9923[0.9825,1.0022]	0.9861[0.9525,1.0209]	0.9533[0.8862,1.0256]
AG+GG	210	0.9999[0.9980,1.0017]	0.9993[0.9945,1.0041]	0.9983[0.9815,1.0154]	0.9999[0.9742,1.0263]
			GCLM		
rs2301022 G/A					
GG	384	Reference	Reference	Reference	Reference
GA	402		0.9973[0.9935,1.0011]	0.9928[0.9787,1.0072]	1.0031[0.9798,1.0269]
AA	142		0.9994[0.9932,1.0057]	0.9926[0.9709,1.0148]	1.0034[0.9704,1.0375]
GA+AA	544	0.9988[0.9973,1.0003]	0.9970[0.9931,1.0008]	0.9928[0.9791,1.0066]	1.0032[0.9812,1.0256]
rs3170633 G/A					
GG	386	Reference	Reference	Reference	Reference
GA	360		1.0032[0.9988,1.0076]	1.0086[0.9936,1.0238]	1.0117[0.9871,1.0369]
AA	184		1.0009[0.9966,1.0052]	1.0016[0.9858,1.0175]	1.0033[0.9755,1.0318]
GA+AA	544	1.0004[0.9989,1.0019]	1.0020[0.9982,1.0058]	1.0022[0.9886,1.0159]	1.0083[0.9863,1.0309]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. Notes: Blank indicates non-converged data.

Table J27: Effects modification of SOD3 genotypes on the association between Cu and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
rs8192287 G/T					
GG	754	Reference	Reference	Reference	Reference
GT	145	1.0000[0.9996,1.0003]	0.9999[0.9989,1.0009]	1.0002[0.9963,1.0041]	1.0013[0.9943,1.0085]
TT	9	1.0008[0.9976,1.0041]	1.0015[0.9935,1.0097]	1.0013[0.9570,1.0476]	0.9914[0.9433,1.0421]
GT+TT	154	1.0000[0.9996,1.0004]	0.9999[0.9989,1.0010]	1.0003[0.9967,1.0040]	1.0021[0.9955,1.0087]
rs699473 C/T					
CC	754	Reference	Reference	Reference	Reference
CT	145	0.9996*[0.9992,0.9999]	0.9989*[0.9979,0.9999]	0.9974[0.9936,1.0012]	0.9953[0.9894,1.0012]
TT	9	0.9997#[0.9993,1.0001]	0.9993[0.9982,1.0003]	0.9981[0.9942,1.0021]	0.9969[0.9903,1.0035]
CT+TT	154	0.9996*[0.9992,1.0000]	0.9991*[0.9981,1.0000]	0.9975[0.9942,1.0008]	0.9971[0.9913,1.0029]
rs13306703 C/T					
CC	606	Reference	Reference	Reference	Reference
CT	264	0.9999[0.9996,1.0002]	0.9997[0.9988,1.0006]	0.9982[0.9950,1.0015]	0.9972[0.9918,1.0027]
TT	39	1.0005[0.9991,1.0019]	1.0004[0.9965,1.0042]	0.9970[0.9815,1.0128]	0.9969[0.9799,1.0142]
CT+TT	303	0.9999[0.9996,1.0002]	0.9997[0.9989,1.0006]	0.9984[0.9952,1.0016]	0.9959[0.9903,1.0015]
rs8192288 G/T					
GG	757	Reference	Reference	Reference	Reference
GT	146	1.0000[0.9996,1.0004]	1.0000[0.9990,1.0010]	1.0011[0.9972,1.0050]	1.0038[0.9971,1.0105]
TT	9	1.001[0.9976,1.0044]	1.0015[0.9935,1.0097]	1.0014[0.9572,1.0477]	0.9969[0.9520,1.0440]
GT+TT	155	1.0000[0.9996,1.0004]	1.0000[0.9990,1.0010]	1.0011[0.9972,1.0051]	1.0038[0.9972,1.0104]

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, *p<0.05.

Table J28: Effects modification of SOD3 genotypes on the association between Ni and annual mean pollutants

Genotype	Subjects (n)	NOx 20m Exp(β) (95% CI)	NO_2 20m Exp(β) (95% CI)	PM_{10} 20m Exp(β) (95% CI)	$PM_{2.5}$ 20m Exp(β) (95% CI)
rs8192287 G/T					
GG	773	Reference	Reference	Reference	Reference
GT	145	0.9982[0.9961,1.0003]	0.9953[0.9898,1.0007]	0.9849[0.9669,1.0032]	
TT	9	0.9893[0.9770,1.0018]	0.9717[0.9433,1.0010]	0.8219**[0.7109,0.9501]	
GT+TT	154	0.9986[0.9966,1.0007]	0.9963[0.9909,1.0018]	0.9843#[0.9666,1.0024]	0.9825[0.9534,1.0125]
rs699473 C/T					
CC	291	Reference	Reference	Reference	Reference
CT	418	0.9989[0.9972,1.0007]	0.9975[0.9929,1.0021]	0.9965[0.9811,1.0122]	0.9867[0.9616,1.0125]
TT	220	1.0002[0.9981,1.0023]	1.0009[0.9956,1.0063]	1.0055[0.9883,1.0230]	0.9997[0.9708,1.0294]
CT+TT	638		0.9987[0.9943,1.0031]	0.9989[0.9839,1.0141]	0.9915[0.9680,1.0156]
rs13306703 C/T					
CC	621	Reference	Reference	Reference	Reference
CT	267	0.9987[0.9970,1.0003]		0.9878[0.9734,1.0024]	0.9861[0.9628,1.0099]
TT	40	0.9959[0.9903,1.0014]		0.9498[0.9003,1.0019]	0.9642[0.9049,1.0273]
CT+TT	307	0.9986#[0.9970,1.0002]		0.9897[0.9753,1.0043]	0.9862[0.9638,1.0090]
rs8192288 G/T					
GG	775	Reference	Reference	Reference	Reference
GT	147	0.9982[0.9961,1.0003]	0.9952[0.9898,1.0007]	0.9845[0.9665,1.0028]	0.9784[0.9493,1.0085]
TT	9	0.9892[0.9769,1.0016]	0.9714[0.9430,1.0007]	0.8198**[0.7097,0.9471]	0.9506[0.7984,1.1319]
GT+TT	156	0.9986[0.9965,1.0008]	0.9963[0.9909,1.0018]	0.9890[0.9709,1.0073]	

Models are adjusted for age, sex, ethnicity, height, BMI, IMD, urinary CCR for smoking exposure and study year including effect for school. #p<0.1, p<0.01.
Notes: Blank indicates non-converged data